UNIVERSITY
OF MICHIGAN

DEC 22 1955

SOCIAL SCIENCE
LIBRARY

Journal of Experimental Psychology

BF
1
J86

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Effect of Psychological Feedback upon Work Decrement: R. B. PAYNE AND G. T. HAUTY 343

A Linear Relationship Between Paired Comparisons and Rank Order: R. T. ROSS 352

Generalization of Reinforcement Among Similar Responses Made in Altered Stimulus Situations: M. H. MARX AND B. B. BERNSTEIN 355

Meaningfulness and Isolation as Factors in Verbal Learning: G. A. KIMBLE AND R. H. DUFORT 361

Frequency of Seeing and Radial Localization of Single and Multiple Visual Stimuli: H. W. LEIBOWITZ, N. A. MYERS, AND D. A. GRANT 369

Distance Judgment by the Method of Fractionation: J. PURDY AND E. J. GIBSON 374

Effects of Complex Transformations of Feedback upon Simple Instrumental Behavior: C. E. NOBLE AND I. G. BROUSSARD 381

Discrimination Decrement as a Function of Time in a Prolonged Vigil: P. BAKAN 387

The Necessary Conditions for Cue-Position Patterning: D. BIRCH AND V. VANDENBERG .. 391

American Psychological Association

Vol. 50 No. 6 December 1955

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1955 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Box 1557, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright © 1955 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 50, NO. 6 DECEMBER, 1955

EFFECT OF PSYCHOLOGICAL FEEDBACK UPON WORK DECREMENT

R. B. PAYNE AND G. T. HAUTY

USAF School of Aviation Medicine

A recent review of the literature by Ammons leaves little doubt that some form of psychological feedback is beneficial to the learning and performance of a wide variety of tasks (1). In fact, this principle has been so thoroughly confirmed in its elementary form that it has become a fundamental consideration in the design of training programs and equipment (4), and its incorporation by general behavior theory poses a major challenge. Nevertheless, one could scarcely say that all the fundamental questions about psychological feedback have been resolved. Optimal understanding and use of the principle in the control of behavior await further work upon such problems as the content, latency, and specificity of the return; the immediacy and endurance of its effects; and the interaction of such effects with proficiency level, work duration, cerebral efficiency, and many other conditions of task performance upon which the return may be imposed.

There is a notable poverty of evidence concerning the effects of feedback upon prolonged performance. Does a supplementary return promote proficiency indefinitely? Does it arrest or forestall work decrement in prolonged repetitive tasks? Opinions on these questions are divided. Bartlett, for example, has said that the most efficient offset to "skill fatigue" so far demonstrated is some means whereby *S* can periodically check his performance against a standard (2). He argues that without such knowledge *S* may unknowingly lower his standard so that his output, although deteriorating, may nevertheless seem acceptable. On the other hand, Hauty and Payne found that *S*s apprised of substandard performance by supplementary auditory or visual cues were unable to forestall work decrement (3). The signals boosted proficiency for the entire period of 7 hr., but they did not affect the rate of decline in any favorable way.

In a work decrement situation it is possible, of course, that the efficacy of feedback will depend upon its relevance to whatever factors are responsible for the deficit. If the decrement is fundamentally a matter of motivational failure, a return having predominantly incentive properties should allay the onset of decline. On the other hand, if the decrement is largely a matter of organizational impairment, a return having predominantly directive properties should

provide an effective countermeasure. In actuality, work decrement almost certainly involves both these factors, as well as a great many others, and this raises the interesting possibility that its arrest might depend upon the additivity of different returns applied concurrently. Such notions cannot be easily tested, for they involve two principal uncertainties—the etiological details of work decrement, and the stimulus functions of the returns. However, it is at least possible to arrange the form and content of returns so that an operational distinction between cue properties and drive properties can be made on logical and intuitive grounds. For example, a very prompt and specific designation of the fact and locus of error might be said to promote proficiency largely because of its directive properties, even though the cues might acquire some drive value. By contrast, a tardy, infrequent signal which merely apprises *S* of substandard performance without suggesting some specific corrective action, if beneficial to performance, would seem so largely because of its incentive properties. Evidence of significant, orthogonal effects attributable to the two types of signals would favor the distinction drawn between them on rational grounds.

The foregoing considerations led to the design of an experiment which explored the single and joint effects of directive and incitive techniques upon work decrement in a complicated compensatory pursuit task. These effects were examined across different levels of cerebral efficiency induced by analeptic and depressant drugs.

Procedure

Subjects.—The *S*s were 144 volunteer, experimentally naive, healthy, rested, basic airmen whose ages ranged narrowly around 20 yr.

Task.—The USAF SAM Multidimensional Pursuit Test (CM813E), described in previous communications (**3, 5**), was used. Briefly, this task involves the scanning of four cam-driven simulated aircraft instruments and the counteraction of their drifts by the coordinated adjustment of simulated aircraft controls. The performance schedule involved a preliminary practice period consisting of 40 1-min. trials, a 10-min. rest period during which experimental factors were introduced, and a final work period consisting of 192 1-min. trials (4 hr.). Score, consisting of the time during which all four pointers were kept on null concurrently, was recorded for each of the five pretreatment cycles and each of the 24 posttreatment cycles (1 cycle = 8 trials).

Experimental conditions.—Directive feedback, which will be called Tuition (T), consisted of three methods of apprising *S* whenever any of the four instrument pointers drifted from its null area. The control level (T_1) consisted of no prescribed warning at all, i.e., the deviations were detectable only through *S*s' own scanning procedure. The second level (T_2) provided, in addition to T_1, a single peripheral visual signal whenever a deviation occurred, but it did not specify the deviant instrument. The third level (T_3) provided, in addition to T_1, individual instrument signals that specified both the fact and the locus of deviation. The supplementary signals were triggered instantaneously by the deviations.

Incitive feedback, which will be called Motivation (M), provided *S*s with varying amounts of information concerning their progress at the conclusion of each cycle of eight trials. The control level (M_1) consisted of no prescribed information at all other than a very general, hazy notion of progress which *S* might form spontaneously. The second level (M_2) provided *S* with knowledge of his standing relative to a performance level represented to him as a group norm. This information was gauged to the nearest $\sigma/4$, and it pertained only to the single cycle then completed. The third level (M_3) included the second, but it also allowed *S*'s standings on previous cycles to remain in view. The relative standings were displayed by differently colored pegs inserted into a wall board drilled with 24 (cycles) $\times$ 24 ($\sigma/4$) holes. The absolute value of the group standard on each cycle was set at 1σ above the cycle mean which had been calculated from normative data, although *S* was never aware of this because the relative standings were always displayed in standard score units taken by *E* from a conversion table. Thus, the group standard represented to *S*s was actually that which would be expected to be equaled or surpassed by about 16% of the parent population.

Cerebral efficiency was varied by means of

TABLE 1

ANALYSIS OF VARIANCE OF ADJUSTED POSTTREATMENT SCORES

Source	*df*	Mean Square	*F*
Tuition	2	5,071,246	78.96*
Motivation	2	669,336	10.42*
Drugs	3	4,032,156	62.78*
T × M	4	116,021	1.81
T × D	6	125,073	1.95
M × D	6	83,790	1.30
T × M × D	12	94,805	1.48
Ss treated alike	108	64,226	
Cycles	23	226,621	116.81*
C × T	46	2,153	1.11
C × M	46	3,522	1.81*
C × D	69	44,140	22.75*
C × T × D	138	3,525	1.82*
C × M × D	138	1,568	<1.00
C × T × M	92	2,273	1.17
C × T × M × D	276	1,901	<1.00
Pooled interaction			
Ss × Cycles	2,484	1,940	
Total	3,455		

* Significant ratios, $p < .005$.

clinical doses of analeptic and depressant drugs (D). The lowest level (D_1) was induced by a mixture of .65 mg. hyoscine hydrobromide with 50 mg. diphenhydramine hydrochloride (Benadryl-hyoscine). Control levels were achieved by no drug at all (D_1) and lactose placebo (D_2). The highest level (D_3) was induced by a 5 mg. dose of dextro-amphetamine sulfate (Dexedrine). All preparations were delivered in No. 1 pink capsules during the interval between preliminary and final practice periods.

The organization of the experiment followed a multiple-factor design consisting of 3 × 3 × 4 = 36 treatment combinations. The serial order of the combinations was randomized across 36 testing days. On any one day four *Ss* were drawn randomly from the population of volunteers, assigned to the combination selected for that day, and tested simultaneously in different rooms under essentially identical conditions.

At the conclusion of the final work period, *Ss* completed a questionnaire designed to gain information about their task attitudes and other dispositions, as reported in a previous investigation (5).

RESULTS AND DISCUSSION

The 3,456 scores (144 *Ss* × 24 cycles) were adjusted for regression upon pretreatment achievement levels. Their variation was then analyzed as shown in Table 1, where Tuition, Motivation, and Drugs constitute the main experimental variables, and Cycles are treated as replications.

Directive feedback (Tuition).—It is apparent from Table 1 that the variation associated with directive feedback is highly significant. A plot of the general means for each treatment, shown in the upper panel of Fig. 1, and the *F* ratio for the C × T interaction indicate that the treatment effects are essentially uniform across all cycles. Tests of significance[1] applied to treatment differences show that T_3 performance is superior to T_2 performance across all cycles ($p < .01$), and that T_2 performance is superior to T_1 performance across all cycles ($p < .01$).

Despite the apparent generality of the directive feedback effect, a further study of Table 1 indicates that the treatment differences are not uniform from one drug group to another throughout all cycles (C × T × D interaction). This state of affairs is elaborated in Fig. 2, which plots the cycle means for each T × D combination. When analyzed by appropriate tests of significance the C × T × D interaction is found due largely to the progressive depreciation of feedback effects under the Benadryl-hyoscine condition ($p \geqq .08$ for 23 of the 24 T × D_4 subclass mean comparisons on the final 8 cycles). To a smaller degree the interaction is due also to the fact that the T_3 effects are not significantly greater than the T_2 effects under the Dexedrine condition beyond the Cycle 16 ($p \geqq .11$ for the eight subclass mean comparisons). The generality of the effect holds, however, for all pharmacological conditions over the first 16 cycles, and

[1] All *t* tests cited in this study were double tailed.

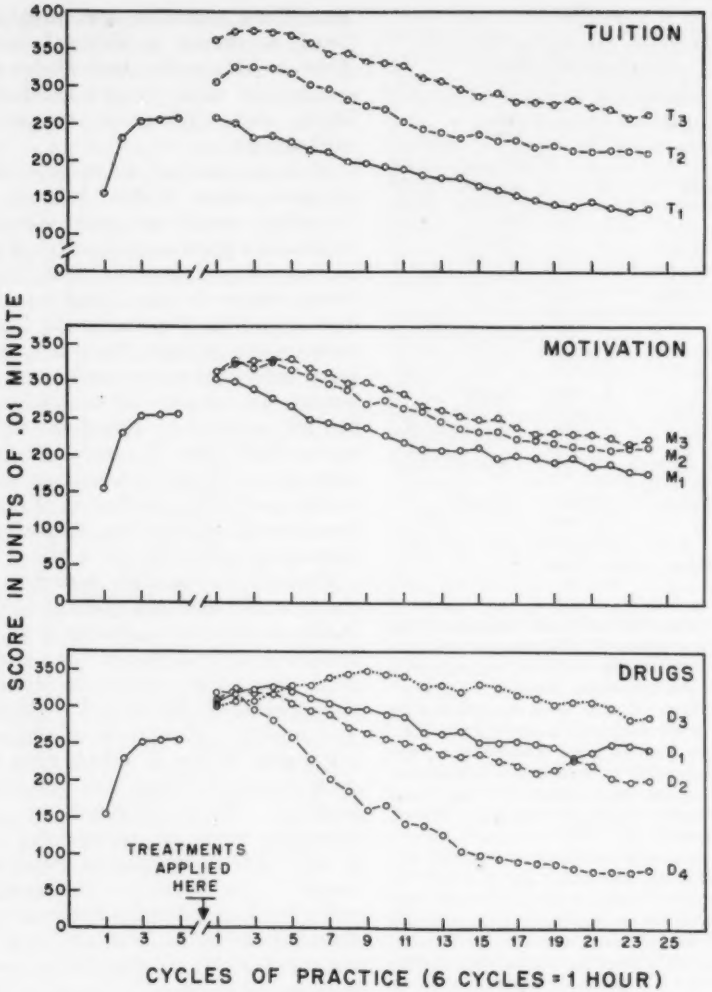


FIG. 1. Effects of principal treatments upon work decrement.

for both the pharmacologically neutral conditions throughout all 24 cycles.

The foregoing evidence shows conclusively that the undrugged *S*'s proficiency is raised uniformly throughout the entire 4 hr. of work by instantaneous warnings of performance error. Since the increment is a positive function of the specificity of the information supplied, it can be ascribed largely to the directive properties of the cues, i.e., the cues impose a more efficient temporal and spatial organization upon *S*'s scanning behavior. Curiously enough, however, the cues failed to offset performance decrement relative to the initial gains which they produced. When *S*'s cerebral efficiency is altered by analeptic or depressant drugs, the stimulus functions

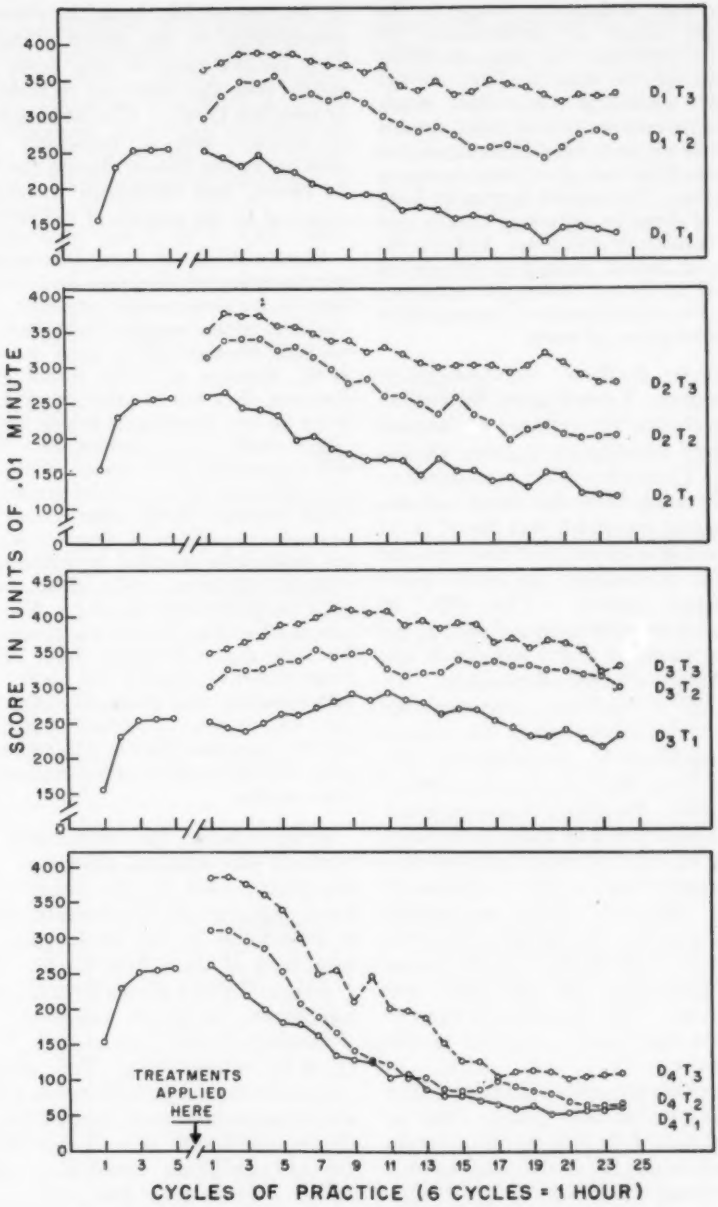


FIG. 2. Directive feedback effects within pharmacological treatments.

of the cues undergo a change in the terminal stages of performance. In cerebral analepsis the locus-specifying function of the cues provides no significant advantage over that which merely indexes the fact of error, a result which is not surprising if analepsis per se allows *S* to work at or near maximum capability. In cerebral depression both types of directive returns gradually lose their regulatory functions, perhaps because of raised excitation thresholds induced by the anticholinergic properties of the Benadryl-hyoscine mixture and by the prolongation of work.

Incitive feedback (Motivation).—Reference to Table 1 again shows that the variation attributable to incitive feedback treatments is very significant. These effects are orthogonal to those arising from the other two experimental variables, but the C × M interaction suggests that they are not uniform throughout the entire post-treatment period. The plot of general treatment means shown in the central panel of Fig. 1 will aid in the visualization of this relationship. An appraisal of treatment differences for each cycle shows that M_3 means are significantly superior to M_1 means from the second cycle onward ($p \leqq .04$). No effect, of course, could have been exerted by these treatments during the first 10 min. because they were applied only at the conclusion of cycles. The M_2 means are significantly superior to the M_1 means for Cycles 2–14 ($p \leqq .05$), but thereafter the differences are generally not significant. The M_3 means are generally not significantly superior to the M_2 means.

Since this type of feedback provided neither specific nor timely cues to what *S* should do in order to maintain his proficiency, its effect could scarcely be attributed to directive properties. Rather, it would seem more reasonably endowed with drive value, arising, in the case of M_2, from the observed relationship to the group standard, and, in the case of M_3, from this source plus the observed relationship to previous levels. The latter source, of course, was a negligible spur relative to the former during the first 14 cycles, but thereafter it became essential to the success of the return.

Certain similarities and differences in the behavior of directive and incitive returns are now readily apparent. Orthogonal to one another, both have immediate positive effects upon performance. Relative to these initial gains, however, they fail to arrest work decrement for any appreciable length of time, either singly or in combination. They are distinguishable in terms of the way in which they behave in the presence of other variables in the experiment. For example, the effects of directive returns are profoundly modified by extremes of cerebral efficiency, while those arising from incitive returns are not. Also, the effects of incitive returns are diminished by task duration, while those arising from directive returns are not. The orthogonality and distinguishability of the two types of return constitute empirical grounds for the different stimulus functions which were imputed to them earlier.

Drugs.—Since the differential effects of the cortical stimulant and depressant used in this study have been demonstrated repeatedly, there is little point to any prolonged consideration of them here (**3, 5**). The *F* ratio in Table 1 shows Drugs to be a significant source of variation, the uniformity of which varies with cycles (C × D interaction). The general treatment means plotted across cycles are shown in the lower panel of Fig. 1. During the first 30 min. of work there are no significant treatment differences. Thereafter the depressant curve drops sharply and remains significantly under all others ($p < .01$),

while the analeptic curve becomes and remains significantly superior to all others ($p < .01$). The C × D interaction, therefore, reflects largely the fact that the performance effects of drugs must await the completion of internal absorption and distribution processes.

Although the C × T × D interaction has already been elaborated in terms of directive feedback treatments within drugs across the cycles, there is some merit to a further examination of it in terms of the differences between drugs. For this purpose the D means within the T subclasses have been plotted across the cycles, as shown in Fig. 3. Ap-

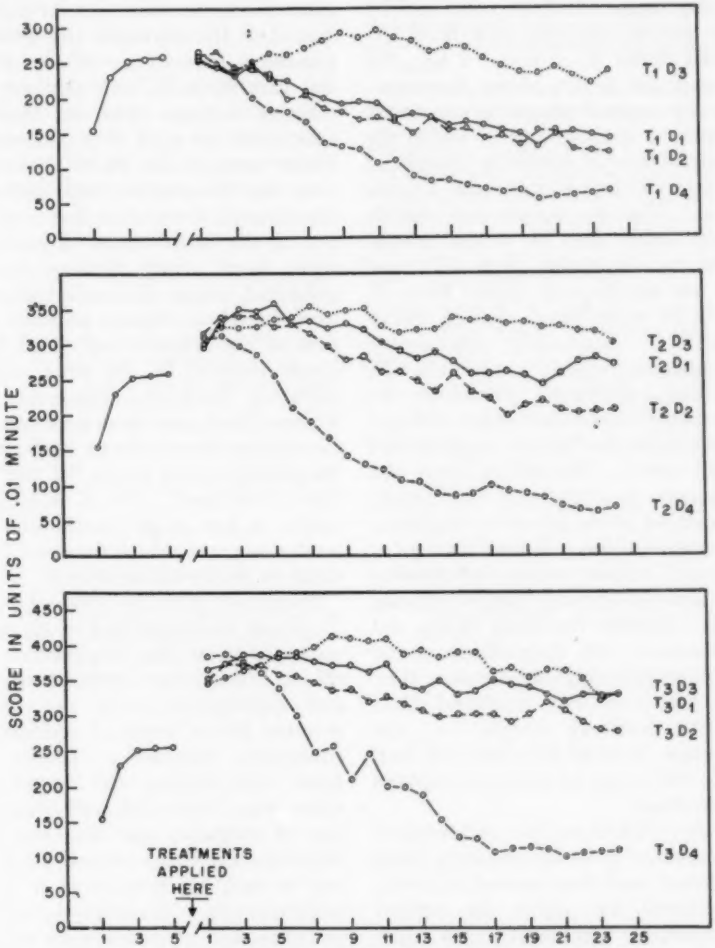


FIG. 3. Pharmacological effects within directive feedback treatments.

propriate tests of significance applied to D differences under the T_1 condition indicate that, after the lapse of absorption-distribution time (approximately 40 min.), the D_3 means are significantly superior to all others and the D_4 means are significantly inferior to all others ($p \leqq .01$). Under the T_2 condition all means are shifted upward during the early cycles, but the D_4 means decline more rapidly than before, reaching the level exhibited under T_1 in about 2 hr. By contrast, the decline of the pharmacologically neutral means is somewhat delayed, in consequence of which the D_3 superiority is neither as prominent nor as prolonged as under the T_1 condition. The D_3 means are significantly higher than D_1 means in only 10 of the 24 cycles ($p \leqq .05$), and they are significantly higher than D_2 means in only 18 of the 24 cycles ($p \leqq .05$). Essentially comparable outcomes are found under the T_3 condition. Evidently, therefore, the extremes of cerebral efficiency induced by the pharmacological preparations placed certain limitations upon the increments in proficiency which could be realized from directive feedback. They seem to have defined a range of capacity within which proficiency variation due to other influences could occur. Incitive feedback effects did not interact with drug effects across the cycles probably only because they were relatively less profound than directive feedback effects, i.e., the increment induced by them fell well within this range of capacity imposed by the drugs.

Cycles.—Changes in performance from cycle to cycle are generally quite significant, and they depend in a very complicated way upon the various experimental treatments and their combinations. The trend of the general means, of course, is downward; but this trend is affected markedly by drugs, and to a smaller but still significant extent by incitive feedback and by D × T combinations. For example, consider the drug influences. Under the depressant condition there is a profound rate of decline during the first 120 min.; and the more the performance is raised initially by directive feedback, the greater the rate of decline (see bottom panel of Fig. 2). By contrast, the analeptic condition induces an *increment* over the first 90 min., and the only significant declines observed later are relative to the peak of this increment. Under none of the $D_3 \times T$ combinations was the output of any later cycle significantly lower than that displayed during the first 20 min. of posttreatment work. Such declines as were exhibited under the analeptic condition, therefore, were a positive function of the singular and joint increments induced by the drug and the directive feedback treatments. It follows from this and previous evidence that the burden of performance sustentation was borne by the analeptic treatment. The C × M interaction is due mainly to the fact that incitive treatments induce an initial delay in performance decline.

Subjective effects and performance.—Evidence was presented in an earlier paper to show that the performance effects induced by cortical analeptics and depressants could not be accounted for in terms of concomitant subjective disposition effects (5). Since this finding was contrary to what many have believed about the role of attitudes and dispositions in performance, there seemed good reason to seek verification of it in the present study. Accordingly, task dispositions and attitudes were sampled by a 13-item questionnaire, similar to the one used earlier, at the conclusion

of the work period. Each *S*'s response to each item was scored on a 3-point scale; and a composite score, which could range from 13 (unfavorable) to 39 (favorable), was derived for each *S*. The variation in terminal performance levels was adjusted for multiple regression upon pretreatment levels and disposition indexes, then analyzed in order to determine the significance of the main experimental effects and their interactions. Suffice it to say that our previous findings with respect to (*a*) the effects of drugs upon subjective dispositions, and (*b*) the functional connection between these and performance effects, were confirmed.

Summary

One hundred forty-four *S*s received preliminary training on a complicated compensatory pursuit task involving simulated aircraft instruments and controls, then continued work for 4 hr. after having been distributed among 36 combinations of three directive feedback treatments, three incitive feedback treatments, and four pharmacological treatments. The treatment effects were appraised in terms of subjective dispositions and task performance, and the functional connection between these outcomes was examined.

Directive feedback signals induced an immediate increment in task proficiency. The increment was a positive function of the specificity of the information supplied, and it remained essentially constant for the duration of the task when all other experimental effects were pooled. However, it was not maintained at former levels during the terminal stages of performance under analeptic and depressant conditions. There was no condition, nor combination of conditions, under which directive feedback could be said to have postponed work decrement.

Incitive returns also induced an immediate increment in task proficiency. This increment was orthogonal to all other experimental effects; but it varied with task duration, first, by growing progressively larger for approximately the first 60 min., then by growing progressively smaller throughout the remainder of the task. Thus, incitive returns could be said to have postponed work decrement for no longer than an hr., either singly or in combination with other experimental variables.

Dextro-amphetamine sulfate was singularly effective in minimizing work decrement, while the mixture of hyoscine and diphenhydramine hydrochloride hastened and maximized it. Significantly divergent task dispositions induced by these preparations were shown to be unable to account for corresponding divergences in performance.

References

1. Ammons, R. B. Knowledge of performance. *USAF, WADC Tech. Rep.*, 1954, No. 54-14.
2. Bartlett, F. C. A note on early signs of skill fatigue. *Great Britain, Air Min., Fly. Pers. Res. Com. Rep.*, 1948, No. 703.
3. Hauty, G. T., & Payne, R. B. Mitigation of work decrement. *J. exp. Psychol.*, 1955, **49**, 60–67.
4. Miller, R. B. Handbook on training and training equipment design. *USAF, WADC Tech. Rep.*, 1953, No. 53–136.
5. Payne, R. B., & Hauty, G. T. The effects of experimentally induced attitudes upon task proficiency. *J. exp. Psychol.*, 1954, **47**, 267–273.

(Received January 13, 1955)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

A LINEAR RELATIONSHIP BETWEEN PAIRED COMPARISONS AND RANK ORDER

ROBERT T. ROSS

Sonoma County Mental Hygiene Clinic, Santa Rosa, California

In 1930, Hevner (**4**) published a study in which scale values arrived at for 20 different samples of handwriting evaulated by 370 *Ss* were compared for the method of paired comparisons, rank order, and equal appearing intervals. She concluded, in part, that *scale values* from the methods of paired comparisons and rank order were linearly related and practically equal.

In order to avoid differences due to the statistical treatment of her data, Hevner resorted to the following procedure:

> The rank order for each of the specimens for each subject was first recorded and especially prepared data sheets were then used to tabulate for each subject the number of times each specimen was preferred to every other specimen. The results of this method, in the final form, are, therefore, comparable to the results of the method of paired comparisons, and the treatment of the results was the same for both methods (4, p. 196).

By comparing the scale values obtained from these derived data with those obtained directly from the method of paired comparisons, Hevner was able to demonstrate the near equivalence of the two sets of scale values. The application of this finding to later work leads to some interesting relationships.

In 1928, Guilford (**1**) had suggested a short-cut method for finding scale values for the method of paired comparisons. This method he later called the Composite-Standard Method and applied it to both paired comparisons and rank-order data (**2**). This procedure deals with the total number of choices for any one item (C) in the method of paired comparisons and with the sum of the ranks ($\sum fR$) for each item in the method of rank order (**3**).

He arrived at proportions from the method of paired comparisons by use of the formula

$$M_p = \frac{C + .5N}{nN}, \qquad (1)$$

where n is the number of items and N the number of judges, and showed that the scale values arrived at by the use of these proportions are linearly related to those obtained from the use of more elaborate procedures (**3**, p. 170).

Similarly, by summing the *rank values* ($\sum fR$) for each item (**3**, p. 179), he arrived at a set of proportions which are treated in the same manner as those from paired comparisons:

$$p_{S>os} = \frac{\sum fR - .5N}{nN}, \qquad (2)$$

where n and N have the same meanings as above and $R = 1$ is the *lowest* rank. (Therefore, if $r = 1$ is the highest rank, $R = n - r + 1$.)

Now since the scale values arrived at by the use of Equations 1 and 2 involve identical statistical procedures, and since Hevner has shown that the scale values arrived at by the two different methods are the same, it follows that the proportions obtained from Equations 1 and 2 must be the same if the same N judges evaluate the same n items. It follows, therefore, that

$$\frac{C + .5N}{nN} = \frac{\sum fR - .5N}{nN}, \qquad (3)$$

whence, if $r = n - R + 1$ (i.e., when $r = 1$ is highest rank),

$$C = Nn - \sum r, \qquad (4)$$

and if we substitute $\sum r = N\bar{r}$, then

$$C = N(n - \bar{r}) \qquad (5)$$

which provides us with a means of predicting what the total number of choices for each item would have been in the method of paired comparisons from a knowledge of the mean ranks obtained by the method of rank order or, conversely, to predict mean ranks from the total number of choices in paired comparisons.

The relationship stated in Equation 5 may be arrived at on purely rational grounds. Suppose all N judges ranked the same item as 1 (most preferred), and suppose that in all paired comparisons they consistently chose this item over all $(n - 1)$ alternatives. Then each judge would choose this item in $(n - 1)$ comparisons and the total number of choices would be

$$C = N(n - 1). \qquad (6)$$

Similarly, if another item were ranked last, or nth, by all N judges, its mean rank would be n. In all of the paired comparisons involving this item, one of the other $(n - 1)$ items would be chosen, so that the total number of choices for this particular item would be zero.

We now have two points of reference as follows

$$\bar{r} = 1, \quad C = N(n - 1)$$
$$\bar{r} = n, \quad C = 0,$$

and if we assume that a straight line joins these points, we arrive at the relationship

$$C = N(n - \bar{r}), \qquad (7)$$

which is the same as Equation 5.

This equation states a relationship which must apply to the raw data of both of the methods and provides the basis for Hevner's demonstration of the equivalence of the scale values obtained from them. It now remains to be shown that the relationship stated in Equations 5 and 7 actually applies to experimental data.

TABLE 1

TOTAL NUMBER OF CHOICES (C) FROM THE METHOD OF PAIRED COMPARISONS AND MEAN RANK ($\bar{r}$) FROM THE METHOD OF RANK ORDER FOR 67 JUDGES ARRANGING 11 ITEMS

S	Obtained C	$\bar{r}$ Derived from C	Obtained $\bar{r}$	C Derived from $\bar{r}$
Sy	447	4.33	4.20	456
GO	421	4.72	4.88	410
LO	448	4.32	4.35	446
SQ	194	8.11	8.07	196
PS	255	7.20	7.02	267
VS	151	8.75	8.71	153
OS	171	8.45	8.13	192
MC	534	3.02	3.67	491
BB	212	7.84	7.67	222
JO	374	5.42	5.43	373
SO	478	3.87	3.82	481

Two sets of empirical data are presented. The first set, shown in Columns 2 and 4 of Table 1, was obtained from 67 university students (40 men and 27 women) who were asked to indicate their preferences for 11 different kinds of music (symphony, grand opera, light opera, string quartet, piano solo, violin solo, organ solo, musical comedy, brass band, jazz orchestra, and swing orchestra) first by the method of paired comparisons and then by the method of rank order. This was accomplished by listing the 55 comparisons on a mimeographed sheet in optimal order (5) in the form

——Symphony Orchestra vs. Jazz Orchestra——
——Light Opera vs. Swing Orchestra——

and requiring that the preferred member of each pair be indicated by a checkmark. The *S*s were then instructed to turn the page and arrange the 11 items in rank order with 1 indicating the most preferred and 11 the least preferred position.

Column 2 of Table 1 shows the total number of choices (C) obtained from the method of paired comparisons and Column 4 shows the mean ranks ($\bar{r}$) obtained from the method of rank order.

These two groups of data are plotted against each other in the upper line of Fig. 1. The line drawn through these points is Equation 5 with $N = 67$ and $n = 11$. (Note that the x-intercept is n, the number of items, and the y-intercept is nN, the denominator of Equations 1 and 2.)

The approximation of the empirical data to the requirements of Equation 5 is sufficiently good to justify accepting the equation as a reasonable statement of the relationship existing between them.

A second set of empirical data is presented in Table 2. By the same procedure as that out-

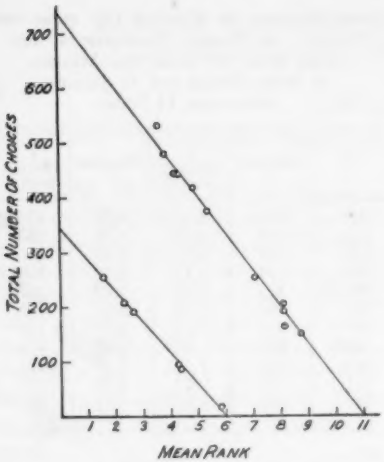


FIG. 1. Total number of choices obtained from the method of paired comparisons plotted against mean rank obtained from the method of rank order. Upper series for 67 judges arranging 11 items; lower series for 57 judges arranging 6 items. Curves are obtained from Equation 5.

lined above, 57 men were asked to express their preference for Masculine Men, Average Men, Effeminate Men, Masculine Women, Average Women, and Effeminate Women, with the difference, however, that a month intervened between the administration of the methods of paired comparisons and rank order. The total number of choices (C) and the mean ranks ($\bar{r}$) are shown in Columns 2 and 4, respectively, of Table 2, and are plotted against each other in the lower line of Fig. 1. The line drawn through the points is Equation 5 with $N = 57$ and $n = 6$. Again the approximation of the theoretical relationship to the raw data is satisfactory.

By the use of Equation 5 it is possible, of course, to transmute total choices to rank orders and vice versa. Columns 3 and 5, respectively, in both tables indicate mean ranks derived from total choices and total choices derived from mean ranks. The closeness of fit of the data to the theoretical relationship may be observed by comparing the obtained and derived values of (C) and ($\bar{r}$) in these two Tables. Equation 4 can be used in this fashion when it is desired to obtain mean ranks from paired comparisons, especially in those cases where scale values are undesirable because of the doubtful normality of the distribution of judgments.

TABLE 2

TOTAL NUMBER OF CHOICES (C) FROM THE METHOD OF PAIRED COMPARISONS AND MEAN RANK ($\bar{r}$) FROM THE METHOD OF RANK ORDER FOR 57 JUDGES ARRANGING SIX ITEMS

n	Obtained C	$\bar{r}$ Derived from C	Obtained $\bar{r}$	C Derived from $\bar{r}$
MM	211	2.30	2.36	207
AM	252	1.58	1.50	255
EM	21	5.63	5.83	10
MW	94	4.35	4.29	97
AW	189	2.68	2.62	193
EW	88	4.46	4.36	94

SUMMARY

It has been shown, both rationally and empirically, than when N judges are asked to indicate their preferences for n items by both the method of paired comparisons and the method of rank order, a linear relationship of the form

$$C = N(n - \bar{r})$$

will be found to hold between the total number of choices (C) from the paired comparisons, and the mean rank ($\bar{r}$) from the rank orders.

Inasmuch as this relationship holds, the proportions and scale values arrived at by the application of Guilford's Composite Standard Method to paired comparisons and rank orders will be identical. The relationship is also implicit in Hevner's procedure by which she demonstrated the equivalence of scale values arrived at from the method of paired comparisons and rank order.

REFERENCES

1. GUILFORD, J. P. The method of paired comparisons as a psychometric method. *Psychol. Rev.*, 1928, **35**, 494–506.
2. GUILFORD, J. P. *Psychometric methods*. New York: McGraw-Hill, 1936.
3. GUILFORD, J. P. *Psychometric methods*. (2nd Ed.) New York: McGraw-Hill, 1954.
4. HEVNER, K. An empirical study of three psychophysical methods. *J. gen. Psychol.*, 1930, **4**, 191–212.
5. ROSS, R. T. Optimal orders in the method of paired comparisons. *J. exp. Psychol.*, 1939, **25**, 414–424.

(Received December 16, 1954)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

GENERALIZATION OF REINFORCEMENT AMONG SIMILAR RESPONSES MADE IN ALTERED STIMULUS SITUATIONS[1]

MELVIN H. MARX AND BENJAMIN B. BERNSTEIN

University of Missouri

Research on the spread of the influence of reward has been confined until recently to the investigation of stimulus-response relationships of the ordinal type described in the typical spread-of-effect design. In a recent study by Marx and Bernstein (**4**) results were obtained, however, which indicated that generalization of reinforcement may occur to stimulus-response connections not in ordinal proximity to reward. It was found that when the response to one of four synonyms in a list of words was rewarded, there occurred a strengthening of the nonrewarded responses made to the three other synonyms independently of their ordinal positions. This result was predicted on the basis of the Marx and Bunch generalization-of-reinforcement hypothesis (**5**), which states that reinforcement of a response has a strengthening influence upon those other responses that are most closely related, along various dimensions, to the reinforced response.

The present experiment was designed to test another implication derived from the generalization-of-reinforcement hypothesis. The specific prediction made was that reward of the selection of one stimulus from a set of differing stimuli would result in an increased tendency toward repetition of a second, nonrewarded selection of one of a number of slightly varied forms of the particular kind of stimulus first selected. In other words, there would be a spread or generalization of reinforcement from the first rewarded response to a nonrewarded response related to it on the basis of stimulus similarity.

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-25631, monitored by the Perceptual and Motor Skills Research Laboratory, Human Resources Research Center. Permission is granted for reproduction, translation, publication, use, and disposition in whole and in part, by or for the United States Government. The opinions or conclusions contained in this report are those of the authors and are not to be construed as reflecting the views or indorsement of the Department of the Air Force. The cooperation of the Army ROTC unit at the University of Missouri in supplying *Ss* for this experiment is gratefully acknowledged.

METHOD

Subjects.—The *Ss* were 210 male college students naive as to the purpose of the experiment. They were obtained from the freshman and sophomore classes of the Army ROTC unit at the University of Missouri. An additional set of 41 similar *Ss* was run in a supplementary experiment to test certain hypotheses suggested as alternative explanations of the major results. The special conditions under which these *Ss* were run are described along with the results.

Apparatus.—An automatic LaBelle 2 × 2-in. slide projector was used to present line drawings of the stimulus material. A Hunter decade timer was used to indicate the time intervals for slide presentation. The *Ss* were seated along two tables which were partitioned by shields so that they could not observe each other's responses. Four *Ss* were placed along a table. Each *S* responded by pressing one of five push buttons on a box that was before him. The buttons were arranged linearly to correspond to five test items similarly projected on the screen, thus permitting *S* to select one of the items by pushing the appropriate button, i.e., the one in the corresponding spatial position. A bank of red lights was on a panel just above the push buttons; all of the lights flashed on to indicate a

correct response. The *Ss*' responses were recorded by two 20-pen Esterline-Angus Operations Recorders, one for each table of *Ss*.

Experimental design.—The general plan of the experiment was to reward each *S* following his selection of one of five stimulus items early in the test series, and then to measure later in the test series the strength of his nonrewarded response to a slide showing five variations of the stimulus item which had been first selected and rewarded. It was predicted that the mean difference between first- and second-trial nonrewarded responses for such experimental *Ss* would be less than the corresponding difference for control *Ss* whose original selection of one of these same five stimulus items was not rewarded.

In order to provide a single first-trial reinforcement for control *Ss* and also to lengthen the test series, two complete sets of five different stimulus items were used. The first set consisted of gauge forms. The *Ss* whose response to one of these was rewarded were in Group GR (gauge rewarded). The second set consisted of island shapes. The *Ss* whose response to one of these was rewarded were in Group IR (island rewarded). IR *Ss* not only provided controls for the GR *Ss* but also, since they were given a comparable island reinforcement, themselves constituted an additional set of experimental *Ss*. Their controls were provided by the GR *Ss*.

A regular graded scale from low to high readings provided the variations within each of the five gauge forms. An increasing amount of the gauge face was blackened to indicate progressively higher readings. Because of the irregularities in the island shapes, no such regular relationship was possible for them. Instead, a focal point, said to represent a military installation, was placed on each island form. This shifted, roughly, in a left to right direction for successive representations of the same island. Since the test-item variations for the gauges were more regular and discriminable, it was predicted that the experimental-control differences would be more marked for Group GR.

Differentiation among the five basic stimulus items of each set was increased by having *Ss* memorize the names for each gauge and island before the test series was presented.

General procedure.—Six to eight *Ss* were trained and tested at one time. Sets of *Ss* were assigned at random to one of the two groups. Preliminary instructions were given to orient *S* to the experimental situation. For purposes of motivation the task was described in terms of military operations. It was explained that military personnel often must remember names of objects and particular details about such objects for communication at some later time. The *Ss* were told that the experiment was designed to test their ability to remember these kinds of information.

Memorization trials.—Study slides were presented with the correct name of the gauge or island beneath each form on the slide. Artificial names were used for islands (Pregas, Fortega, Kialo, Rokoj, Ankau). Common gauge names were used for gauges (Gas, Oil, Heat, Time, Speed).

Following this, the separate items were presented one at a time with all five names. Each *S* was to choose the correct name for each item by pushing the button beneath it. If *S* made a correct choice, the bank of red lights flashed on. This complete cycle of study slides followed by the separate item slides was presented four times for both islands and gauges.

Test trials.—Following the training task, *Ss* were given instructions for the test trials. As an introduction they were told that occasions arose in which more detailed information had to be remembered, such as the *location* of an army base on an island, or a particular *reading* on a gauge. It was emphasized that now they had to learn the correct answers as they went along, starting from no information and gradually building up knowledge about the slides as they were given "right" and "wrong" information following their responses. It was explained that the right answers had been assigned at random, and not according to any set pattern or system.

A trial consisted of 12 slides, each remaining on for 8 sec. For each slide *S* had to make a choice from five possible answers. As before, the red lights were used to inform *S* when his answer was the correct one.

Each slide projected line drawings of five objects. The *S*'s task was to pick out the particular object he thought might be the correct one and respond by pressing the appropriate button. Two kinds of slides were employed: (*a*) *Master slides:* These showed all of the five gauges, or islands, which were the same objects on which the *Ss* were trained. The first two slides of each trial were master slides. (*b*) *Detail slides:* These 10 slides showed each gauge, or island, in the five variations of its basic form (five degrees of gauge blackening, or five varying focal points on islands).

The island and gauge slides were presented alternately. For Group GR, the island master slide preceded the gauge master slide, and island detail slides preceded gauge detail slides in pairs throughout the list. For Group IR the reverse sequence was used. The ordering of the 10 detail slides was systematically rotated for different sets of *Ss* to control for position effects. Four trials were given by repeating the same list of 12 slides. A gray rest slide was presented for 16 sec. between trials. The *Ss* were told that their score would be the total number of correct answers they gave.

Only a single reinforcement was given in each of the first two trials. The reinforcement was

administered to any response made on the *second* master slide in the list. The last two trials were not analyzed but were given with two additional reinforcements in order to avoid arousing the suspicion of the *S*s.

Analysis of data.—Strength of response was measured by the numerical difference between the push-button responses made to the same slide on the first and second trials (cf. **3, 4, 6**). The push buttons were considered to be numbered from 1 to 5, from left to right across the response box. A strong response was therefore one which was minimally varied in ordinal number units.

Thus, if a response of 3 were given in the first trial, a response of 2 or 4 to the same stimulus on the second trial (represented by a response-difference score of 1) was considered as evidence of greater strength than a response of 1 or 5 on the second trial (represented by a response-difference score of 2). A repetition of the response originally given, of course, represented maximum strength. The response variation score was in that case zero. Response-variation scores of increasing magnitude thus indicated decreasing degrees of response strength.

In the present study a response-difference score was obtained for each *S* from the first- and second-trial responses made to the slide presenting the five readings on the gauge which had been chosen, and reinforced, on the second master slide. The same kind of measure was obtained when the response to an island shape had been reinforced. This measure will be referred to as an E (experimental) score. A difference score for each *S* was also obtained from the responses to that slide presenting the stimulus which had been chosen, with responses not reinforced, on the first master slide. This measure will be referred to as the C (control) score. Each *S* thus served as an experimental *S* for one category and a control *S* for the other.[2]

The data for a number of *S*s were eliminated from the major analyses for gross defects produced by failure to follow instructions. Obvious sequences and too many or too few responses were the common flaws. Thirty-seven of the original 149 *S*s from Group GR and 26 of the 124 *S*s in Group IR were eliminated. The χ^2 for the difference in proportion of these *S*s was 1.59 (3.84 required at the .05 level).

In order to obtain a more unambiguous test of the hypothesis, records were also eliminated if *S* failed to repeat the reinforced master-slide response on the second trial. These could not easily be used, since the critical detail slide could not be determined when two *different* master slides had been reinforced. Twenty-four *S*s were eliminated from Group GR and 20 from Group IR for thus failing to repeat.

[2] An exception to this generalization was made for seven *S*s of the 210 who provided usable records. These seven *S*s could not be used for control purposes because they failed to make the first-trial first-slide response which was used to determine the critical detail slide for control scoring. Their records were otherwise satisfactory and so they were not discarded.

TABLE 1

MEAN INTERTRIAL RESPONSE-DIFFERENCE SCORES

Group		*N*	Mean	*SD*	*t*
GR	E	112	1.55	1.11	2.69**
	C	93	1.98	1.18	
IR	E	98	1.70	1.27	.65
	C	110	1.81	1.12	
GR + *IR*	E	210	1.62	1.18	2.45*
	C	203†	1.89	1.15	

* $P = .02$.
** $P = .01$.
† See footnote 2 for an explanation of this discrepancy in the control *N*.

RESULTS

Major E/C analyses.—The mean response-difference scores for the critical detail slide are shown in Table 1. The mean E score for Group GR was significantly lower, indicating greater response strength, than the mean C score. The difference for Group IR was also in the predicted direction but was not statistically significant. This more clear-cut E/C difference for the GR Group had been predicted on the basis of the greater regularity of the response scale on the critical detail slide for that group.

The distributions of scores were not skewed enough to distort results obtained from the *t* or *F* test. The variances did not differ significantly.

Combination of Groups GR and IR gave a mean E/C difference which was significant at the .02 level. Since the within-*S*s correlation between E and C scores in both groups was not significant and the groups to be combined did not differ significantly, the unmatched *t* test was applied. Al-

though a precise reference to parent population cannot be made, the combinations of groups are presented to provide some picture of the over-all effect of the experimental variable. Any question concerning these combinations does not apply to the individual group *t* tests.

If the E scores tended to distribute differently from the C scores on the first trial, a chance distribution of responses on the second trial might have resulted in differences between group means. Therefore, the distributions of first-trial responses for the two groups were plotted. The distributions were all essentially rectangular. The difference scores to be expected, given a chance rectangular distribution on the second trial, were calculated for the four groups. The range was only from 1.56 to 1.58. This negligible difference indicates that the obtained mean differences presented in Table 1 cannot be attributed to any differences in the distributions of first-trial responses.

Exact repetition of response was also used as a measure of response strength. Both groups showed nonsignificant differences in the predicted direction (16% to 9% exact repetitions for GR, and 18% to 10% for IR).

The data from *Ss* who failed to repeat the rewarded master slide were checked. There was a consistently greater proportion of exact error repetition in the eliminated *Ss*. The mean response-difference scores were also lower. These results suggest that these *Ss* constituted a somewhat different population, perhaps inferior in motivation and/or ability, since they avoided their errors to a lesser degree. In any case, all the E/C differences, although small, were in the predicted direction.

Control of intratrial generalization effects.—If there was a closer relationship in the experimental *S* on each trial between the reinforced master-slide response and the erroneous response made on the critical detail slide, then the apparent intertrial response strengthening could well be an artifact of intratrial generalization effects. This generalization from master slide to detail slide could have occurred along two dimensions. First, the degree of blackening of the various gauges tended to differentiate them to varying degrees from the wholly unblackened master-slide gauges. Differentiation of the island shapes by positioning of the focal point did not produce this kind of gradient. Second, the spatial relations among the five push buttons offered a linear physical dimension for response generalization for both GR and IR *Ss*.

Frequency-distribution analysis of the GR data provided little support for generalization along the blackening dimension. There was actually a tendency for fewer *Ss* to choose the least blackened gauges: *n*'s of 12, 21, 27, 29, and 23, from least to most blackened gauges. This compares with the control distribution of 19, 19, 14, 26, and 15 *Ss*.

The experimental GR *Ss* did show evidence of intratrial response generalization along the push-button dimension. Differences were computed between the master-slide button response and the subsequent critical detail-slide button response. The difference between the E and C distributions of difference scores was significant at just short of the .05 level ($\chi^2 = 7.46$ for 3 *df*; 7.82 required). The experimental IR *Ss* showed a similar trend, but their distribution was not significantly different from that of their controls ($\chi^2 = 3.34$ for 3 *df*).

Stronger intertrial response-difference mean scores were found for the *Ss* with least intratrial variation:

1.41, 1.40, 1.71, and 2.08, for 0, 1, 2, and 3–4 difference classes. No gradient was found for their controls: 2.13, 1.70, 2.11, and 2.32, for comparable positions. At each point the experimental mean was stronger.

Partial correlations between first- and second-trial detail-slide responses are presented in Table 2. There was a significant intertrial error correlation with the relationship between reinforced-response button position and critical detail-slide error button position partialled out. This result indicates that this kind of intratrial generalization was not sufficient to account for the intertrial strengthening.

One further check was made. The 12 experimental GR *Ss* who had selected the most similar (least blackened) of the differentiated gauges were compared with the 19 control *Ss* who selected this gauge. These experimental *Ss* showed an unexpectedly low intertrial mean difference score (1.00) compared with the high control score (2.58). The obtained difference was highly significant ($t = 3.51$, $P < .01$). Both mean scores should have been high, since the least blackened gauge corresponded to Button 1. Intertrial differences as high as 4 were possible when the response on the first trial was to Buttons 1 or 5. The comparable IR mean scores were both high (1.81 for 21 experimental *Ss*, 1.91 for 23 controls).

When these 12 experimental *Ss* and the 19 comparable controls were eliminated from the analyses, both the response-difference measure and the partial-correlation measure still favored the experimental GR *Ss*, but failed of statistical significance.

TABLE 2

INTERTRIAL CORRELATIONS FOR RESPONSES ON CRITICAL DETAIL SLIDE WITH RELATIONSHIP TO REINFORCED RESPONSE PARTIALLED OUT

Group		*N*	r	$r_{12.RR}$	*CR*
Main Experiment					
GR	E	112	.06	.05	2.69**
	C	93	−.27	−.27	
IR	E	98	−.05	−.05	.71
	C	110	−.15	−.15	
Supplementary Experiment					
GR	E	20	.59	.62	2.10*
	C	21	−.03	−.03	
IR	E	21	.28	.24	—
	C	20	.23	.30	

* $P = .05$.
** $P = .02$.

Supplementary experiment.—The relatively low intertrial correlations for experimental *Ss* (cf. Table 2) indicate that with only 12 stimulus slides in the series the well-motivated *Ss* were able to remember and thus to avoid errors. This suggests an alternative explanation of the results. It is that the errors were simply less well remembered and thus less well avoided in the experimental groups. This hypothesis assumes that there was no true strengthening of an S-R relationship in the experimental group, but rather a weakening of retention of the errors.

To test this hypothesis, the main experiment first trial was duplicated with more ROTC *Ss*. On the second trial, *Ss* were told to reproduce, as exactly as possible, all their first-trial responses. It was emphasized that the new instructions applied to errors as well as to correct responses. This follows the general procedure introduced by Nuttin (**7**).

Superior rather than inferior retention was found for the experimental condition. The mean intertrial difference score for the experimental critical detail slide for the 20 GR *Ss* who repeated the rewarded response was .65, compared with 1.14 for their 21 controls. The corresponding value for the 21 experimental IR *Ss* was .95, compared with 1.45 for the 20 controls.

It may be assumed that a certain amount of guesswork was involved in the attempt to reproduce each of the first-trial errors. In this process any differential first-trial strengthening of

error tendencies as a function of reinforcement would make for a greater tendency towards repetition even under the new instructions. Strong statistical support for the intertrial strengthening interpretation is provided by the greater partial *r* for GR *Ss*, shown in the lower part of Table 2, since this controlled for any influence of intratrial generalization mediated along the push-button dimension. Moreover, this result cannot be attributed to the influence of intratrial generalization along the blackening dimension. Only two GR *Ss* selected the least blackened gauge on the first-trial critical detail slide, and their mean intertrial difference score (1.00) was above the group mean (.65).

Discussion

The fact that the stimulus situation was changed from master slide to detail slide makes any simple application of response generalization insufficient as an explanation of the intertrial relationships. Stimulus generalization is similarly inadequate, since the E/C difference in number of exact repetitions was not significant. However, Hull's suggestion of a combination of stimulus *and* response generalization (**1**, p. 183) could explain the results. Such S-R generalization clearly overlaps with the generalization-of-reinforcement principle, particularly if stimulus mediation is explicitly involved, as in the present design.

However interpreted, the present results point up the need for analysis of complex stimulus and response generalization effects resulting from reinforcement. No empirical investigation has previously been directed toward the S-R generalization suggestion made by Hull. Moreover, this suggestion has received little theoretical attention, or even *ad hoc* utilization. As a matter of fact, it is not explicitly dealt with in Hull's latest work (**2**).

Summary

An experiment designed to test the Marx and Bunch (**5**) generalization-of-reinforcement hypothesis is reported. The *Ss* were first trained to discriminate among related stimulus items presented by a slide projector. Their response to one of five such forms was then reinforced, and the strength of a response to five variations of the reinforced stimulus was later measured. A reciprocal experimental-control design was used whereby each *S* served both as an experimental *S* for one set of stimulus items and a control for the other.

Significantly stronger experimental responses occurred for the stimulus items (gauge forms) for which a more regular scale of stimulus variation was used. The second set of items (island shapes) also showed a difference in the predicted direction, but it was short of statistical significance.

Control analyses suggested that the major experimental-control differences could not be explained as artifacts of intratrial generalization. A control experiment showed that the results were not due to a weakening of retention of errors due to the experimental treatment.

The results lend some further support to the generalization-of-reinforcement hypothesis, or to the Hullian suggestion of stimulus-response generalization (**1**). Taken in conjunction with the positive results previously obtained (**4**), they suggest that spread or generalization of the influence of reinforcement need not be restricted to relations of ordinal proximity.

References

1. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
2. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
3. Marx, M. H. Error-reinforcement gradients with bias in the response series controlled. *Amer. Psychologist,* 1951, **6,** 293. (Abstract)
4. Marx, M. H., & Bernstein, B. Generalization of reinforcement among responses to synonyms. *J. gen. Psychol.,* 1955, **52,** 49–64.
5. Marx, M. H., & Bunch, M. E. New gradients of error reinforcement in multiple-choice human learning. *J. exp. Psychol.,* 1951, **41,** 93–104.
6. Marx, M. H., & Goodson, F. E. Further gradients of error reinforcement following repeated reinforced responses. *Amer. Psychologist,* 1952, **7,** 274–275. (Abstract)
7. Nuttin, J. "Spread" in recalling failure and success. *J. exp. Psychol.,* 1949, **39,** 690–699.

(Received November 15, 1954)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

MEANINGFULNESS AND ISOLATION AS FACTORS IN VERBAL LEARNING

GREGORY A. KIMBLE AND ROBERT H. DUFORT[1]

Duke University

This series of experiments began as a study of the relationship between speed of learning and meaningfulness, with the intention of sampling this function for a wider range of meaningfulness values than had previously been employed. The results of the first experiment, however, suggested very strongly that the relationship could not be studied in an uncomplicated way using the procedures employed there, because of an unanticipated effect of meaningfulness. Experiments II and III examine this complicating factor more closely.

Experiment I

The purpose of Exp. I was to obtain functions relating speed of learning to meaningfulness, using the paired-associate learning procedure. This method has the advantage of making it possible to include in a single set of learning materials items of varying degrees of meaningfulness, so that the complete function can be investigated on a single group of *S*s. The experiment was performed in two replications which will be referred to as IA and IB. Except for the learning materials, they were identical.

Method

Learning materials.—For each of the two replications, ten pairs of terms were selected from Noble's (**4**) list on the basis of the following criteria: (*a*) They covered the complete range of *m* (meaningfulness) values. (*b*) Only words of nearly the same *m* value were paired. (*c*) Except for two pairs of paralogs (pronounceable nonsense words) of nearly the same *m* value, an attempt was made to obtain a series of roughly equal steps along the meaningfulness continuum. (*d*) The lists for the two replications were different. The three words in common were used to make it possible to sample the entire meaningfulness continuum in each list. The pairs of terms and their *m* values appear in Table 1.

These pairs of words were typed on memory-drum tapes and presented to *S*s in six different haphazard orders on different trials. If *S* did not learn the list in six trials, the order in which the pairs were presented in Trials 7–12 was the same as that on Trials 1–6. Only two *S*s in Exp. IA and one in Exp. IB required more than 12 trials. For these *S*s previous orders were repeated a second time. There was, however, no evidence that any *S*s learned the order of the response terms.

Subjects and procedure.—The *S*s in each replication were 20 Duke University students. They received these instructions: "This is not a personality test of any sort. I am interested only in the words, not in the reactions of the subject. You will be shown a series of ten pairs of words. The first time the list is shown, you are to try to learn as many of the pairs as possible. Do not try to learn the order in which the items appear. After the list has been presented once, it will be shown again: the same pairs but in different order. When you see the stimulus word, reply with the word paired with it in the first list. If you are not sure of the correct response, guess at it. The experiment will be continued until you get the entire list correct once." The *S*s were not told that some of the items were not real words, nor were they told the purpose of the experiment.

The words were presented by a Gerbrands memory drum which exposed the stimulus term for 3 sec. alone, and then presented the stimulus and response terms for 3 sec. together. The 3-sec. exposure was selected on the basis of preliminary investigation as one which produced a minimum of difficulty in deciding whether the response was made soon enough to be counted as correct.

[1] We should like to thank Mr. James Hall for collecting the data for Exp. IA, and Mr. P. J. Baugh and Mr. William Rudge for running Exp. III.

TABLE 1

PAIRS OF WORDS AND THE *m*-VALUE OF EACH

Experiment IA				Experiment IB			
Stimulus	*m* value	Response	*m* value	Stimulus	*m* value	Response	*m* value
NEGLAN	1.04	MEARDON	1.05	GOJEY	0.99	MEARDON	1.05
TAROP	1.24	XYLEM	1.24	BYSSUS	1.13	NEGLAN	1.04
FEMUR	2.09	LOZENGE	2.09	FLOTSAM	2.19	GRAPNEL	2.15
PALLOR	3.06	SEQUENCE	3.21	RAMPART	3.36	ARGON	3.34
ORDEAL	3.91	ZENITH	4.44	QUOTA	4.68	YEOMAN	4.60
QUARRY	5.10	EFFORT	5.13	MALLET	5.61	KENNEL	5.52
QUARTER	5.98	REGION	5.98	UNCLE	6.57	ZERO	6.15
ZEBRA	7.12	GARMENT	7.17	JEWEL	7.58	INSECT	7.39
WAGON	8.12	DINNER	8.33	OFFICE	7.95	HEAVEN	7.91
ARMY	9.43	KITCHEN	9.61	MONEY	8.98	KITCHEN	9.61

Results and Discussion

The results appear in Fig. 1 which presents median number of trials to the first correct association for each pair of words, for the two replications. The abscissa values are the average meaningfulness values of the two words in each pair. Medians are used as measures of learning because the distributions of the response measures were systematically skewed. The two curves are similar, although that obtained in Exp. IB is less regular than that obtained in Exp. IA. By comparison with the results of previous experiments on the relationship between speed of learning and meaningfulness, these curves differ chiefly at the low-meaningfulness end of the continuum. Earlier investigators (**3, 5**) have obtained functions relating speed of learning to meaningfulness which were negatively accelerated and decreasing. In these previous experiments, nonsense materials were much more difficult to learn than any meaningful materials. The curves in Fig. 1, by contrast, show at most a slightly heightened difficulty of learning one of the nonsense pairs. The others were learned more easily than several of the pairs of meaningful words. Presumably this difference between the present results and those of previous studies is somehow assignable to the paired-associate procedure. Noble's (**5**) results which were based on the same materials as those used here, but employing the serial learning technique, resemble the results of other serial learning studies more nearly than they do these. This suggests that the paired-associate procedure introduces a variable which makes the less meaningful materials easy to learn.

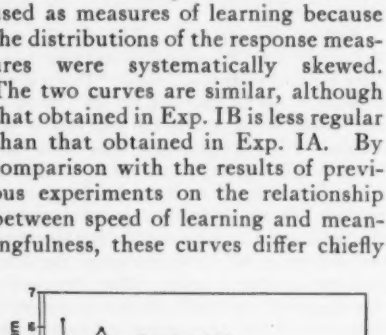


FIG. 1. Median number of trials to learn each pair of terms in Exp. IA and IB. The abscissa values are mean *m* value of each pair.

One factor which might possibly act in this way is the isolation effect studied by von Restorff (6). She found that an isolated term such as a two-place number in a list of nonsense syllables was learned very easily. Her interpretation was in

terms of the figure-ground relationship stressed by gestalt psychologists. According to this explanation, materials which are different from the context in which they are embedded stand out from that background and are learned easily. Just why such isolation should lead to easy learning is a matter which we shall discuss in connection with Exp. III. For our present purposes, the implication of this line of reasoning is that relatively meaningless materials stand out against the otherwise meaningful materials in lists of paired associates and that this isolation or contrast effect aids learning. Experiments II and III are a test of this general idea.

Experiment II

The suggestion that a process like the one investigated by von Restorff is responsible for the precise form of the functions obtained in Exp. I implies the operation of a perceptual factor in rote learning. The purpose of this experiment was to test this implication. Specifically, it seemed likely that such a perceptual process would have a greater influence on the stimulus terms in paired-associate learning than on the response terms. Such reasoning leads to the conclusion that nonsense words in a paired-associate learning experiment would be easy to learn if these words were placed on the stimulus side, but that this effect would not appear when the same words are on the response side of the pair.

Method

Learning materials.—The list of paired associates learned by two different groups of *Ss* in this experiment appear in Table 2. It will be observed that each pair consists of a common three-letter word and a word selected from Noble's list. Of the latter, the two terms, *Xylem* and *Tarop*, are paralogs and would presumably be meaningless or nearly so to all *Ss*. The remaining words are ones which almost all college students would at least recognize. The three-letter words are all of the consonant-vowel-consonant variety, and the vowel is always *a*.

TABLE 2

Pairs of Words Used in Exp. II

Scaled Word	Paired Word	Scaled Word	Paired Word
TAROP	MAN	QUARRY	HAT
XYLEM	RAG	REGION	HAM
FEMUR	TAX	ZEBRA	WAR
PALLOR	CAT	WAGON	JAB
ORDEAL	PAW	ARMY	BAY

This restriction upon the form of the words was based upon a preliminary experiment which was exactly like this one, except that the three-letter words were a heterogeneous collection without this systematic similarity. The results of the preliminary experiment were also like those of this one, but there was an unsatisfactorily large degree of variability in the data. The construction of this list represents an attempt to reduce such variability.

In terms of the lists presented in Table 2, the prediction to which the hypothesis under examination leads is that the two nonsense words, *Tarop* and *Xylem*, would be relatively easier to learn when they appear as stimulus words than when they are response words. (This prediction is on the assumption that the isolation phenomenon suggested by the results of Exp. I is a perceptual effect which would affect the stimulus terms more than the response terms.)

Subjects and procedure.—The *Ss* were 28 Duke University students; 20 of them had served in Exp. IB. Since the results of this experiment were very similar to those obtained from naive *Ss* in the preliminary experiment, it seems unlikely that their previous practice at learning paired associates had any serious effect upon the outcome of this experiment.

One group of 14 *Ss* learned the materials in Table 2 under an arrangement where the scaled words were the stimulus words. The other group of 14 *Ss* learned them with the scaled words as the response words. These two groups will be designated, respectively, as Group S (scaled words on the stimulus side), and Group R (scaled words on the response side).

All procedural details were the same as in Exp. I. Since there were two groups in this experiment, care was taken to be sure that the order of pairs was the same on each trial for each group. Thus any differences in learning must be attributed to the placement of the scaled words on the stimulus or on the response side of the list.

Results and Discussion

The hypothesis under investigation in this experiment was that the opera-

tion of a perceptual factor in paired-associate learning would make nonsense words relatively easy to learn when they were used as stimulus terms, but that no such benefit would appear when the same words were used as response terms. The results, in terms of number of trials to learn, confirm this prediction. For Group S, the mean number of trials to learn the two pairs containing meaningless words as stimuli was 2.00; the mean number of trials to learn the pairs with meaningful words as stimuli was 2.36. For Group R, the mean number of trials required to learn the comparable pairs were 3.57 and 2.59. In this group, of course, the words referred to above as stimuli were on the response side of the pair. On the assumption that most of the scaled words employed in this experiment are less meaningful than the three-letter words, a further implication of these findings is that the effect of meaningfulness on learning may involve a response process as well as the stimulus one investigated directly in this experiment. The evidence for this is that both kinds of words were mastered more quickly by *S*s in Group S.

Both of these interpretations are supported by the results of an analysis of variance presented in Table 3. The significant *F* for experimental conditions supports the incidental observation suggesting an effect of meaningfulness through the response process. For the main purpose of the experiment, the critical *F* test is that for the interaction variance. The fact that it is significant shows that the difference between the number of trials to learn the two kinds of pairs was different from group to group. The sort of difference involved may be summarized by pointing out that, for Group S, the paralogs were easier to learn than the words; for Group R the words were easier to learn than the paralogs.

TABLE 3

ANALYSIS OF VARIANCE ON RESULTS OF EXP. II

Source	*df*	Mean Square	*F*	*p*
Conditions	1	11.52	10.87	<.005
Error between	26	1.06	—	—
Words	1	1.32	1.80	>.20
Interaction	1	6.17	8.45	<.01
Error within	26	.73	—	—

The relevance of these results to the interpretation given for those of Exp. I is that they indicate that the isolating effect from which terms with little meaning benefit occurs only if these terms are used as stimuli. In this sense they support the implication that the isolating effect is a perceptual one which would presumably operate chiefly on stimuli.

EXPERIMENT III

Perhaps the best-known effect of isolation on verbal learning is its influence on the serial position function. When an item different in kind from those constituting the rest of the serial list is inserted into the middle of the list, the usual bow-shaped form of the serial position curve is altered and a dip appears at the point where the isolated term occurs. The assumption that meaningfulness (really the lack of it) can serve to isolate a term leads to an interesting experimental prediction. If a nonsense term is put into the middle of a list of otherwise meaningful words, it should be learned easily. Furthermore, if degree of meaningfulness determines the degree of isolation, it should be possible to manipulate the magnitude of this effect by varying the degree of meaningfulness of the critical word. The interesting feature of this pre-

diction is that, in such an experiment, the speed of learning should vary inversely with degree of meaningfulness. The least meaningful term should be the easiest to learn. This, of course, is exactly the opposite of the outcome of most experiments on the relationship between meaningfulness and learning. Experiment III tests this prediction for four degrees of meaningfulness.

Method

Learning materials.—Four groups of *Ss* learned serial lists of 13 words which were identical except for the middle word. For each group, 12 of the words were 12 of Noble's 14 most meaningful words. The order of these 12 words in the list was: *heaven, dinner, captain, kitchen, garment, jewel, village, wagon, army, zebra, money, insect.* For the four groups, different words were inserted at the seventh position, following *jewel.* These words with their associated *m* values were: *gojey* (0.99), *rampart* (3.22), *kennel* (5.52), and *office* (7.95). The last word, *office,* is from the same meaningfulness range as the other 12 words. The others came from less meaningful portions of the continuum. *Gojey* is a paralog.

Subjects and procedure.—The *Ss* were 80 Duke University undergraduates, in four groups, with 15 men and 5 women in each group. Several additional *Ss* were discarded on the basis of two criteria: (*a*) They reported learning the list in a narrative fashion, and (*b*) they learned the list in five trials or fewer.

Instructions to *Ss* were Yale (2) instructions modified chiefly by omitting the instruction not to make up fanciful associations among the words. Preliminary work indicated that this instruction produced exactly the opposite of the desired effect.

Each *S* learned the list to one errorless repetition. The words were exposed by a Gerbrands memory drum at a 3-sec. rate. Lists (trials) were separated by 5 sec.

Results

The purpose of this experiment was to determine the form of serial position curve for lists which are identical except for the middle word and, with the same exception, fairly homogeneous with respect to the meaningfulness of the individual items. One such presentation of the results appears in Fig. 2. The plot there is in terms of the mean number of errors at each position in the series. Figure 3 presents the same relationship in terms of mean number of *intrusion* errors. Each of these presentations shows the dip predicted for the curve representing the data for the list containing the nonsense word, *gojey,* and each shows that the word, *rampart,* which is next lowest on the meaningfulness continuum was next easiest to learn. In terms of total number of errors, *kennel* was slightly harder to learn than *office* which represents an inversion of the order predicted. In terms of intrusion errors, however, there is a regular increase from *gojey* which produced the fewest to *office* which produced the most. Analysis of variance reveals

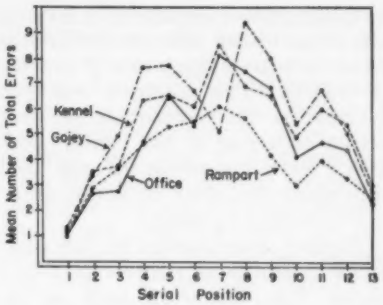


FIG. 2. Mean number of total errors as a function of serial position.

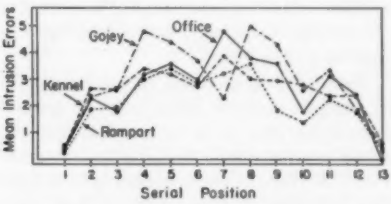


FIG. 3. Mean number of intrusion errors as a function of serial position.

that the relative difficulty of learning the critical word differed significantly from group to group ($F = 9.01$; $p < .001$). This analysis was performed on the percentage of total errors occurring in the critical position. Such treatment of the data minimizes the effect of initial differences in learning ability.

Further examination of the functions presented in either Fig. 2 or 3 shows that the word used as the middle term apparently had some effect on the ease with which several of the immediately preceding and following words were learned. This influence is most obvious for the list containing *gojey* in which the adjacent words show a marked rise in number of errors. Detailed inspection of the records of individual *S*s provided no satisfactory explanation for such an effect.

Discussion

The results of these three experiments leave little room for doubt that an isolation factor benefits the learning of nonsense materials when they are presented among more meaningful materials. In Exp. I this effect produced a flattening of the curve at the low-meaningfulness end of the continuum. In Exp. II it was demonstrated that the effect is chiefly on the stimulus terms, which is compatible with the idea that the process is a perceptual one. Experiment III provided additional support for such an interpretation by demonstrating, as predicted, that the ease of learning the middle word in a list of highly meaningful words varies inversely with the meaningfulness of that term.

Results such as those just summarized fit readily into the framework of gestalt theorizing and pose a problem for the more conventional S-R point of view. For S-R theory, the problem is to reduce the isolation phenomenon to S-R concepts. Gibson (1) has proposed a theory which will do this for the results of Exp. I and II, which involved paired-associate learning. With some additions the same ideas will handle the results of Exp. III. Gibson's approach to paired-associate learning treats the situation as one in which the stimulus terms in the list have the same properties as other stimuli and the response terms have the same properties as other responses. On this general assumption, Gibson then proceeds to the hypothesis that the response terms will sometimes be evoked by stimulus terms other than those to which they are correct as a result of stimulus generalization. To the extent that the various stimulus terms in a paired-associate list are similar, every stimulus word will tend to attract, as responses, words from other parts of the list. Such responses will sometimes occur as overt errors; at other times they will merely inhibit the correct response which also tends to occur. The difficulty of learning any pair of words will, by this reasoning, vary directly with the number of competing associations the stimulus word attracts. Since the process responsible for such attraction is stimulus generalization, stimulus terms which differ markedly from those in the rest of the list will attract few associations and be learned with relative ease. To account in these terms for the ease with which the relatively meaningless items were learned in Exp. I and II, it is necessary to make two assumptions: (*a*) That meaningfulness is an attribute that applies to the stimulus aspect of words, so that it can provide a dimension along which stimulus generalization can operate. The procedures used for scaling meaningfulness and the results of Exp. II both support this assumption. (*b*) That the terms with very low degrees of meaning differ more on this dimension from several of the more meaningful terms than these latter terms differ among themselves. Then it follows that paralogs, for example, will receive relatively few competing associations from other items in the list, and will be learned rapidly.

The extension of such reasoning to the results of Exp. III poses an interesting

problem. In that experiment, both total number of errors and number of intrusion errors at the critical point in the list were smallest when the correct response was a nonsense word, and greater when the correct response was a meaningful word. The results, thus, confirm those of the previous two experiments in showing that, under certain circumstances, an item with little meaning can be learned easily. The problem is one of interpretation. In the previous experiments, it was possible to attribute the differences in ease of learning to differences in the various stimulus words involved in different pairs of terms. In this experiment, however, the stimulus word was the same for each group, the word *jewel*. The problem is to explain how, in the different groups, the word *jewel* could have attracted different numbers of associations leading to different numbers of errors, and how the number of such associations could have been determined by the meaningfulness of a response word which did not occur.

About the only plausible solution to this problem appears to be in the assumption that the correct response did actually occur in some implicit or covert form and that, functioning as a stimulus, it called up associations which interfered with its own appearance. The data provide evidence in favor of part of this assumption in the fact that the results of Exp. III are much the same whether they are expressed in terms of intrusion errors or in terms of total errors. This supports the contention that the results are explainable in terms of competing associations.

There remains the problem of accounting for the dependence of such competing associations upon the meaningfulness of the correct response. Two possibilities suggest themselves: (*a*) A meaningful word like *kitchen*, by the very definition of meaningfulness, possesses more associative bonds than a meaningless word like *gojey*. It might be that these previously acquired associations are called up and that they interfere with the overt appearance of the correct response. (*b*) Following Gibson's hypothesis, a meaningful word like *kitchen* would be expected to attract more generalized associations from within the list because of its similarity to other items. Such evidence as comes from the experiment supports the latter interpretation. About 98% of the intrusion errors obtained from *S*s in Exp. III were responses which would have been correct at some other point in the list. If the associations which *S* brings to the experiment with him were important, more extraneous intrusion errors should have occurred.

Summary

Three experiments on the relationship between meaningfulness and speed of learning are reported.

In Exp. I, 40 college students learned paired-associate lists in which different pairs of terms were of different degrees of meaningfulness. The results of this experiment differed from those obtained in previous experiments chiefly in that the terms with least meaning were much easier to learn than would have been predicted from earlier research. This suggested the operation of the isolation factor studied by the gestalt psychologists.

Experiment II was performed to determine whether the isolation effect would be greater when the isolated nonsense terms were stimulus terms, as compared with response terms. Two groups of 14 *S*s learned paired-associate lists in which one term was a very common three-letter word and the other was a disyllabic word from Noble's (4) list. Of the latter, two were paralogs. One group learned this list with the disyllables as stimulus words; for the other group the three-letter common words were the stimulus words. Results of the experiment confirmed the prediction from the isolation hypothesis that the disyllabic paralogs would be easier to learn as stimuli than as responses.

In Exp. III, four groups of 20 *S*s learned lists of words which were identical except for the middle term. The middle term varied in meaningfulness for the four different groups. This manipulation had a marked effect upon the form of the serial position curve. The list which contained a nonsense term in the middle resembled that usually obtained in studies of the isolation phenomenon, showing a dip in the middle. For the other three groups, the serial position curves were more nearly "normal," but in general, the more meaningful the middle term, the harder it was to learn.

An S-R interpretation patterned after Gibson's (1) theory is offered for the results of these experiments.

References

1. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
2. Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., & Fitch, F. B. *Mathematico-deductive theory of rote learning*. New Haven: Yale Univ. Press, 1940.
3. McGeoch, J. A. The influence of associative value upon the difficulty of nonsense-syllable lists. *J. genet. Psychol.*, 1930, **37**, 421–426.
4. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
5. Noble, C. E. The role of stimulus meaning (*m*) in serial verbal learning. *J. exp. Psychol.*, 1952, **43**, 437–446.
6. von Restorff, H. Über die Wirkung von Bereichsbildungen im Spurenfeld. In W. Köhler and H. von Restorff, Analyse von Vorgängen im Spurenfeld. I. *Psychol. Forsch.*, 1933, **18**, 299–342.

(Received December 6, 1954)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

FREQUENCY OF SEEING AND RADIAL LOCALIZATION OF SINGLE AND MULTIPLE VISUAL STIMULI [1]

H. W. LEIBOWITZ, NANCY A. MYERS, AND D. A. GRANT [2]
University of Wisconsin

In a recent experiment (1) the accuracy of radial localization of a single visual stimulus was investigated as a function of the luminance and the duration of exposure. The results demonstrated that radial localization accuracy of a single stimulus, when visible, is constant for all luminance-duration combinations. Increase in either the luminance or duration of exposure, while increasing the frequency of seeing the stimulus, had no concomitant effect upon the accuracy of localization. The results were interpreted as demonstrating that the magnitude of the primary sensory events does not limit the accuracy of the mechanism of localization, since the variables of luminance and duration, which have a well-known relation to such primary processes (**2**–**5**), have no effect upon localization accuracy. The present experiment extends the original study to determine the effect upon localization accuracy, under various conditions of luminance and duration of exposure, of multiple as well as of single stimuli.

Apparatus and Procedure

The basic instrument employed was a modified Dodge tachistoscope [3] which provides for the rapid alternation of adapting and stimulus fields for controlled durations of exposure. The adapting field was dark with a fixation point in the center, and was replaced for durations ranging from .01 to .64 sec. with the stimulus field consisting of one or more orange, circular stimuli, provided by neon glow tubes, viewed against a black background. The tubes were mounted in holes drilled on a board, and the side facing *S* was covered with onionskin paper. At the completion of the desired stimulus exposure, the adapting field reappeared immediately and, .25 sec. later, radial white grid lines, each 5.85° long and spaced at 10° intervals, appeared in the adapting field for approximately 1 sec. The *S*s observed monocularly through a 2.5 diameter artificial pupil at a viewing distance of 24.5 in. The stimuli, the centers of which are always 4.2° from the center of the field, could be presented in any desired radial position. Duration of exposure was adjusted by means of a mechanical timer in the tachistoscope lighting circuit, while luminance was controlled by placing neutral filters in front of the artificial pupil. The maximum luminance of the stimuli, when there were no filters in the apparatus, was 5.12 mL. The *S*s, male students who were paid for their services, were asked after 10 min. dark adaptation to observe a fixation point in the center of the field and to report in which of 72 radial positions the stimulus or stimuli had appeared. The stimuli were presented randomly in all 72 positions at fixed luminance-exposure time combinations. The *S*s estimated the positions of the targets within 5° in terms of a system of polar coordinates which had previously been explained to them.

Two experiments were performed. In Exp. I, a fixed number of stimuli (i.e., 2, 3, 4, 5, or 6) were presented to five groups of six *S*s each. In Exp. II, six *S*s received random presentations of stimuli varying in number from one through seven. The number of luminance-duration combinations employed was determined by the time required to test *S*s. In Exp. I, for the 2, 3, and 4 stimulus groups, luminance values of .02, .08, .32, 1.28, and 5.12 mL were used in conjunction with durations of .01, .04, .16, and .64 sec., making a total of 20 luminance-duration combinations. For the 5 and 6 stimulus groups, luminance values of .32 and 5.12 mL were used, together with durations of .01, .08, and .64 sec., yielding six combinations. All group members

[1] This research was supported in part by the United States Air Force under Contract AF 33(038)-23294, monitored by the Armament Systems Personnel Research Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Denver, Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] Presented at the annual research conference on vision, Ohio State University, 1954. The authors wish to acknowledge the assistance of J. F. Lomont in obtaining the data.

[3] Manufactured by Ralph Gerbrands, Arlington, Mass.

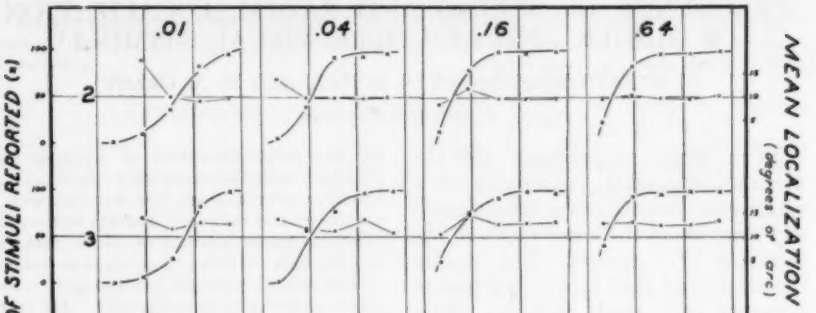


FIG. 1. Frequency of seeing functions and mean radial localization error per reported stimulus as a function of luminance at exposure durations of .01, .04, .16, and .64 sec. for presentations of 2, 3, or 4 stimuli.

received 50 sets of stimuli under each luminance-duration combination. Thus, each member of the 2, 3, and 4 stimulus groups received 1,000 sets of stimuli; those in the 5 and 6 stimulus groups saw 300 sets of stimuli. In Exp. II, only two combinations were employed, .32 mL at .01 sec. and 5.12 mL at .64 sec. exposure. Fifty sets of each of the varied numbered stimuli (1 through 7) were presented at each of these combinations, giving a total of 700 sets of stimuli presented to these *S*s.

The experimental design provided for the random variation of the position of the stimuli, as well as randomization over *S*s of the order in which the luminance-duration combinations were presented. Analysis was made of the "frequency of seeing" of the stimuli, the localization error, and possible learning effects.

RESULTS

Effects Due to Practice

No learning effects were observed in any phase of the present study. With adequate instructions, the *S*'s performance accuracy did not vary with practice.

Frequency of seeing and localization accuracy.—In order to demonstrate the relationship between frequency of seeing and the localization error, both sets of data for the 2, 3, and 4 stimulus cases are plotted in Fig. 1.[4] This graph illustrates the lack of correspondence previously observed for a single stimulus (1), between the frequency with which the stimuli are seen, and the radial localization error. The previous observation, that localization accuracy is unaffected by an increase in the stimulus energy, is here confirmed for 2, 3, and 4 stimuli. This principle also applies to the 5 and 6 stimulus cases which were investigated for a more limited range of luminance values.[4]

Localization error as a function of the number of stimuli.—It will be observed in Fig. 1 that although the localization error is constant for a given lumin-

[4] Tables from which Fig. 1 and 2 were constructed have been deposited with the American Documentation Institute. Order document No. 4685 remitting $1.25 for microfilm or $1.25 for photocopies.

ance-duration combination, the magnitude of the error increases with the number of stimuli. This relation is demonstrated in Fig. 2 where the average localization error is averaged over all luminance-duration combinations and plotted as a function of the number of stimuli employed for both Exp. I and II. The function is similar for Exp. I in which six different *Ss* served in each stimulus-number group, and for Exp. II in which the same *Ss* were presented with varying numbers of stimuli. (The datum for the one stimulus case for Exp. I has been taken from the first paper of this series [1].) In both cases, localization error increases with increase in the number of stimuli, rapidly at first and then more slowly, before approaching a limiting value.

Percentage of stimuli reported and percentage of stimuli reported correctly. —In the present study the *Ss* are asked to estimate the radial position of the stimuli under various luminance-duration conditions and for varying numbers of presented stimuli. At the same time, they are reporting the number of stimuli seen, a response measure which is of interest in connection with the formulation by Schlosberg (6) to provide a rationale for the observation by Hunter and Sigler (5) that the stimulus energy required to perceive *correctly* a given number of stimuli increases with the number presented. Schlosberg has assumed that the perception of a single stimulus is an independent event, so that given the probability of seeing any one stimulus as P, the probability of seeing all N stimuli which are presented will be P^n. This assumption of independence would predict that the percentage of stimuli reported will be independent of the number presented, provided the number is less than the limit of the "span of perception" (5). Table 1, which presents the percentage of stimuli reported and the percentage of stimuli reported correctly may be considered in the light of Schlosberg's assumptions. At the 5.12 mL—.64-sec. combination, the data are clearly not in accord with the hypotheses, although the role of eye movements at this longer duration may introduce a complicating factor. The percentage of stimuli reported and the percentage of stimuli reported correctly are plotted in Fig. 3 for the .32 mL—.01-sec. combination. The percentage of stimuli reported is in agreement with the assumption of independence. Schlosberg's theoretical curve, computed on the basis of the observed probability of seeing one stimulus, has been fitted to the data for the percentage of stimuli reported correctly. The agreement is good for the 1, 2, 3, and 4 stimulus cases, but the theoretical function overestimates the data for 5, 6, and 7 stimuli.

Discussion

The results of the present experiments confirm the previous conclusion concerning the independence of the

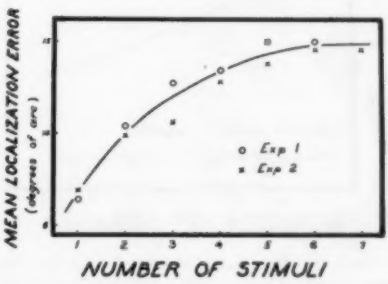


Fig. 2. Mean localization error per reported stimulus in degrees of arc as a function of the number of stimuli presented. Exp. I: the number of stimuli was known and constant for a given subject. Exp. II: the number of stimuli was variable for all subjects.

TABLE 1

PERCENTAGE OF STIMULI REPORTED AND PERCENTAGE OF STIMULI REPORTED CORRECTLY AS A FUNCTION OF THE NUMBER OF STIMULI PRESENTED IN EXP. II AT TWO LUMINANCE-DURATION COMBINATIONS

Luminance-duration Combination	Number of Stimuli Presented						
	1	2	3	4	5	6	7
Percentage of Stimuli Reported							
.32 mL-.01 sec.	90.3	88.0	92.7	88.2	88.2	85.2	82.1
5.12 mL-.64 sec.	100	99.3	98.7	94.6	94.8	92.2	90.7
Percentage of Stimuli Reported Correctly							
.32 mL-.01 sec.	87.7	79.3	76.7	64.3	44.0	34.0	34.7
5.12 mL-.64 sec.	99.3	98.7	93.7	92.3	79.7	65.7	55.3

magnitude of the primary sensory events, which are controlled by the variables of intensity and time, from the mechanism of localization which is not affected by experimental manipulation of these same variables (1). This conclusion is seen to be valid for multiple as well as single stimuli. The data suggest that the localization takes place some time after the primary sensory events are completed, and that if the primary excitation is great enough to produce a visible stimulus, the localization process follows with a degree of accuracy which is independent of the magnitude of the primary excitation.

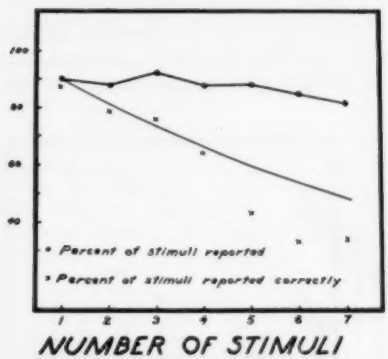


FIG. 3. Percentage of stimuli reported (o) and percentage of stimuli reported correctly (x) as a function of the number of stimuli presented. The theoretical curve fitted to the number of stimuli reported correctly is based on the assumption that the perception of N stimuli will be the n^{th} power of the probability of seeing one stimulus.

The increase in the mean localization error as a function of the number of stimuli presented was found to hold both for Exp. I, in which the number of stimuli was constant and known to *S*s, and for Exp. II, in which the stimulus number was variable.[4] In either procedure, accuracy of localization per visible stimulus decreases with an increase in the number of stimulus elements presented.

SUMMARY

1. Single and multiple stimuli were presented on the circumference of an imaginary circle for various luminance-exposure duration combinations. The *S*s estimated the radial position of the stimulus or stimuli which varied in number from one through seven.

2. Radial localization accuracy is independent of the variables of lumi-

nance and duration, provided the stimulus or stimuli are visible.

3. An increase in the number of stimuli presented increases the radial localization error per stimulus. The error function increases rapidly at first with an increase in the number of stimuli, and then more slowly, before approaching a limiting value.

4. The assumption by Schlosberg that the perception of a stimulus is an independent event was found to hold for a .32 mL—.01-sec. luminance-duration combination but not for a 5.12 mL—.64-sec. combination. In the former case, the assumption that the probability of detecting *exactly N* stimuli is the n^{th} power of the probability of seeing one stimulus was in good agreement with the data for 2, 3, and 4, stimuli but overestimated the results for the 5, 6, and 7 stimuli.

5. The mechanism of localization is not limited by the magnitude of the primary excitatory processes, the latter being completed before the localization mechanism is operative. No explanation is offered for the observed increase in the localization error per stimulus as a function of the number of stimuli presented.

References

1. Leibowitz, H., Myers, N. A., & Grant, D. A. Radial localization of a single stimulus as a function of luminance and duration of exposure. *J. Opt. Soc. Amer.*, 1955, **45**, 76–78.
2. Hartline, H. K. Intensity and duration in the excitation of single photoreceptor units. *J. cell. comp. Physiol.*, 1934, **5**, 229–247.
3. Graham, C. H., & Margaria, R. Area and the intensity-time relation in the peripheral retina. *Amer. J. Physiol.*, 1935, **113**, 299–305.
4. Graham, C. H., & Kemp, E. H. Brightness discrimination as a function of the duration of the increment in intensity. *J. gen. Physiol.*, 1938, **21**, 635–650.
5. Hunter, W. S., & Sigler, M. The span of visual discrimination as a function of time and intensity of stimulation. *J. exp. Psychol.*, 1940, **26**, 160–179.
6. Schlosberg, H. A probability formulation of the Hunter-Sigler effect. *J. exp. Psychol.*, 1948, **38**, 155–167.

(Received January 1, 1955)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

DISTANCE JUDGMENT BY THE METHOD OF FRACTIONATION [1]

JEAN PURDY AND ELEANOR J. GIBSON

Cornell University

The literature of space perception abounds in both experiments and theoretical discussions of the perceived size and distance of objects. The continuum of distance itself, however, has seldom been specifically studied as a psychological dimension. That distance must be perceived accurately for objects to remain constant in their dimensions is a generally accepted proposition. Errors of constancy, indeed, have been attributed to errors in the estimation of distance (**3**; **5**, p. 151). It is time to inquire, therefore, what kind of psychological scale for a dimension of continuing distance exists. Can an observer tell when far and near stretches of distance are equal, or when one is half or twice the other? Gibson predicted that constancy would hold for the distances between objects (**2**, p. 165) on the hypothesis that constancy is a property not merely of perceived objects but of perceived tridimensional space itself. Little evidence is available on this question. Gilinsky (**3**, p. 473 ff.) had two *O*s bisect distances from 8 to 200 ft. long, starting from *O*. She predicted that physically equal units of distance would appear shorter as their distance from *O* increased. This result apparently was found in her data, but the small number of *O*'s used make further experiments desirable, especially since the results of recent size constancy experiments indicate that constancy or even "overconstancy" (**4**) holds for objects at very long distances. It would appear reasonable in the light of such a result, that "constancy," or better, objective accuracy, should characterize the judgments of *all* the physical dimensions of a surface like the ground.

In the present experiment the nature of a psychological scale for distance was investigated by a method of fractionation. The *O* was asked to bisect or trisect distances along an imaginary line stretching over the ground from *O* to an indicated marker. A further group of *O*s was available whose judgments were corrected as they were made. These fractionation judgments had served as a training procedure in another experiment (**1**).

METHOD

The experiment took place in a quadrangle 350 yd. long. The ground was mowed grass, intersected by three walks. Trees and buildings lined the sides, but the view to the end of the quadrangle was unobstructed. A 300-yd. stretch was chosen for *O*'s scaling operations. The *O* stood at one end of the quadrangle, accompanied by *E*. A bicycle and rider moved up or down the quadrangle, as needed, as an indicator of the division point which *O* was asked to find. The bicycle rider had a prearranged schedule of movements to follow and could be signalled by a whistle when *O* wanted him to stop. The area had been surveyed and had marks at 25-yd. intervals to guide the rider, but these marks were invisible to *O*.

Each *O* made 10 fractionation judgments. Six of these were bisections and four were trisections. A "bisection" is the act of stopping the bicycle at half the distance covered by the stretch; a "trisection" is the act of stopping the bicycle at one-third the distance covered, the

[1] This research was supported in part by the United States Air Force under Contract No. 33(038)-22373 monitored by the Perceptual and Motor Skills Laboratory, Human Resources Research Center, Lackland Air Force Base. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

TABLE 1

PLAN OF FRACTIONATION JUDGMENTS

Serial No. of Judgment	Bicycle Approaching		Bicycle Withdrawing	
	Stretches Bisected (yd.)	Stretches Trisected (yd.)	Stretches Bisected (yd.)	Stretches Trisected (yd.)
1	300 (300 to 0)		300 (0 to 300)	
2	150 (150 to 0)		150 (150 to 300)	
3	150 (300 to 150)		150 (0 to 150)	
4		75 (300 to 225)		75 (0 to 75)
5		75 (225 to 150)	50 (25 to 75)	
6	50 (200 to 150)			75 (75 to 150)
7		75 (150 to 75)	50 (100 to 150)	
8	50 (125 to 75)			75 (150 to 225)
9		75 (75 to 0)	50 (175 to 225)	
10	50 (50 to 0)			75 (225 to 300)

judgment in the latter case being the first third or the nearest third of the stretch covered. Table 1 shows the actual stretches divided, and the order in which the fractionations were made. The plan was such that the true division points for the 10 judgments made up a series at 25-yd. intervals from 25 through 250 yd. as indicated in the left-hand column of Table 2. The *O*s were divided into two equal subgroups depending on whether the bicycle rode away from *O* or toward him. The order of judging, and whether *O* halved or trisected, varied between the two subgroups so as to keep within bounds the mileage required of the bicycle rider. The first three bisections, however, were comparable for withdrawal and approach trials.

Before *O* was asked to make a given judgment, markers were set up to indicate the two ends of the stretch that he was to divide. The markers were plain white rectangles, unlabeled. In some trials, *O*'s own station point was the origin of the distance stretch indicated, so that only one marker was required.

The *O* was brought to his station point, given his instructions, and asked to turn his back while the markers were set up. When these were in place and the bicycle at the starting point of the stretch to be divided, *E* asked *O* to turn and judge when the moving target had reached the division point of the specified fraction. The *E* signalled the rider when *O* said "Stop." If *O* was not satisfied with his first setting, one further adjustment was permitted. He then faced the other way during the interval between judgments. The error was measured by the bicycle rider with the aid of a fixed reference point that was invisible to *O*. There were 40 *O*s, half with the bicycle approaching and half with the bicycle withdrawing.

The separate group of *O*s who were corrected made the same judgments, except that *O* was permitted to see the direction and extent of his error, and permanent yard markers labeled with the number of yards distant were set up at the appropriate division points as the judgments were completed. There were 67 *O*s in this group, of whom 53 watched the bicycle approach and 14 watched it withdraw. All *O*s were Airmen in basic training at Sampson Air Force Base.

RESULTS

The data obtained from the bisections and trisections are presented in Table 2. Group mean constant errors (CE) in yards and *SD*'s are given for each of the 10 true division points of the stretches divided. A positive error means that *O* divided the stretch so that the segment nearer him was too large; a negative error indicates that *O* divided the stretch so that the segment nearer him was too small. (It may be noted that if the subject divided the stretch into optical extents of equal visual angle, the nearer segment, by optical geometry, would be much too small.) The length of the interval fractionated and its terminal points can be found by consulting Table 1.

Accuracy of fractionation.—The first question concerns the relation of the judged half-way or third-of-the-way point to the actual division point. Figure 1 presents these data for the

TABLE 2
GROUP MEAN CE AND *SD* OF FRACTIONATION JUDGMENTS

Distance of True Division Point, Yd.	Without Correction						With Correction					
	Target Approaching			Target Withdrawing			Target Approaching			Target Withdrawing		
	Serial No. of Judgment	Mean CE yd.	*SD* yd.	Serial No. of Judgment	Mean CE yd.	*SD* yd.	Serial No. of Judgment	Mean CE yd.	*SD* yd.	Serial No. of Judgment	Mean CE yd.	*SD* yd.
25	10	−1.2*	2.0	4	−2.8*	2.3	10	+2.3*	3.7	4	+4.1*	3.6
50	9	−1.1	3.7	5	+0.7	2.6	9	−0.5	4.6	5	−1.7	3.3
75	2	+5.9*	11.1	3	+0.4	13.7	2	+2.1	8.7	3	+0.1	3.1
100	8	+0.5	3.8	6	+1.6	3.9	8	+1.5*	3.4	6	+3.1*	4.1
125	7	−0.3	3.7	7	+1.8*	3.4	7	−1.1	4.8	7	+5.7*	7.8
150	1	+21.6*	19.0	1	−2.8	23.9	1	+17.5*	19.4	1	+7.6	16.9
175	6	−2.0*	4.3	8	−0.4	6.1	6	−0.1	4.6	8	+2.4*	2.7
200	5	−1.0	8.1	9	−2.9*	4.5	5	−7.4*	11.8	9	−12.4*	6.4
225	3	+6.6*	12.6	2	−5.8	16.2	3	+6.4*	15.1	2	−12.4*	17.0
250	4	+1.6	9.8	10	−4.7*	7.4	4	+6.2*	10.4	10	−0.9	11.9

* CE significantly different from 0 ($p < .05$).

main group of *O*s who were not given knowledge of their errors. The most striking feature of these results is the small amount of error for many of the judgments. Only for one act of fractionation out of 20 does the mean CE go above 6½ yd. When *O* indicates one-half or one-third of a distance stretch, the fraction corresponds very well with the measured half or third. Nine of the 20 CE's are significantly different from zero (see Table 2, uncorrected group), but even the largest is only 7.2% of the total distance fractionated. On the average, the CE is 3.1% of the total distance fractionated. There is a tendency for the CE to be large when

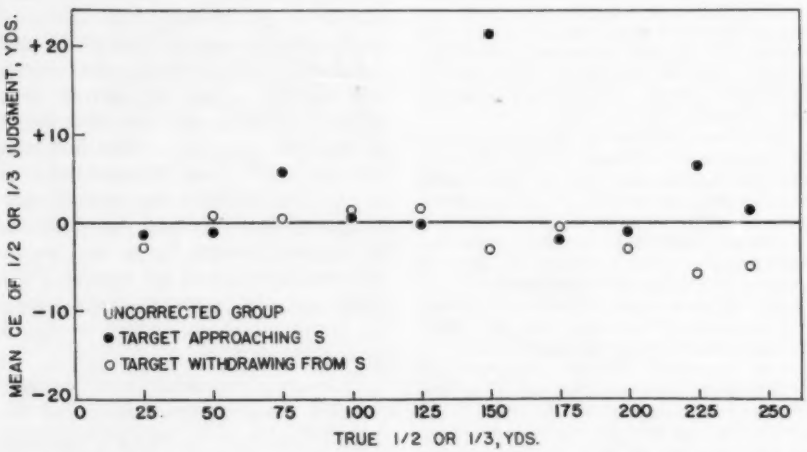


FIG. 1. Plot of the group mean constant error of the judged fraction (in yd. from the true point) relative to the true point of division, arranged in a continuum of distance.

large distance stretches are fractionated. The average error (without regard to sign) is directly related to the total distance fractionated, increasing as the length of the stretch to be fractionated increases. It ranges from 3.2 for a 50-yd. bisection to 21.2 for the 300-yd. bisection.

The size and direction of the error appear to be affected by the direction of movement of the "target." When the two uncorrected subgroups are compared in the first three judgments, the means differ significantly at 150 and 225 yd. The means for approach and withdrawal also differ on fractionations represented by points at 250 yd. and 25 yd. However, these last two differences are not necessarily due to the direction of movement of the target alone. Other differences confounded with direction of movement for all judgments after the third were: (*a*) the judgments were made in different orders, (*b*) *O*s were not always making the same kinds of fractionation of the same distance stretch (see Table 1). Thus the difference between approach and withdrawal CE's for fractionations at 25 and 250 yd. could have been due to any of these variations in procedure. In general, the error appears to be positive on approach trials but much less positive or negative on withdrawal trials. At the 150-yd. division point, for instance, 88% of the approach group but only 53% of the withdrawal group made positive errors.

No difference in the variability of fractionation was produced by the direction of target movement and the kind of accompanying fractionation except at 200 yd. where the *SD* for the approach condition was higher.

A comparison of the accuracy of bisections and trisections can be made by comparing the mean CE's of halving 50-yd. stretches with the mean CE's of taking one-third of 75-yd. stretches. By such a comparison, the size of the CE is no larger for division into thirds than for division into halves, though group variability is larger for the one-third judgments.

Variability was also influenced by the distance from *O* of the stretch to be fractionated. Trisections of 75 yd. can be compared for four stretches which begin at 0, 75, 150, and 225 yds. from *O*. As the distance of the stretch from *O* increases, the *SD* also increases.

Indications of distance "constancy." —It might be supposed that perceptual judgments of distance magnitudes are based wholly or in part on the visual angles subtended at the retina by the stretches of distance. If two physically equal stretches, one near and one far, were compared on this basis, the resulting relative subjective magnitudes would not be equal; the farther stretch would subtend a smaller visual angle and would appear less. In order to make it *appear* equal, the further stretch would have to be actually greater (a negative error in our experiment). Thus if *O* made fractionation judgments wholly or partly on the basis of such information, a negative CE would be the result. Gilinsky (**3**) may have assumed something like this in stating that "Perceived distances are foreshortened. The perceived distance *d* increases with the true distance *D* but at a reduced and diminishing rate" (p. 462). The present data, however, do not substantiate such a hypothesis. The errors in general are positive—that is, *O* made the nearer segment too large in comparison with the farther. And the high degree of accuracy in general indicates that distance judgment exhibits a kind of constancy analogous

to size judgments. The term *length or depth constancy*, as distinguished from *width or height constancy*, might be applied to *O*'s judgments. The fact is that a distance stretch beginning at his own feet can be matched fairly accurately with another stretch beginning at a distance from his feet along the same axis.

In Fig. 2 the mean distance judged one-half has been plotted for the three distance stretches which began at *O*'s feet: 50, 150, and 300 yd. Judgments for approach and withdrawal trials are presented separately, as well as the means for the two combined. The 45° line describes correct half judgments. Though points for the approach condition diverge from accuracy, the mean values lie extremely close to that line. There is a slight indication that as the distance to be halved increases, *O*'s subjective half-way point may move farther away from him so that the nearer segment is physically larger than the farther segment.

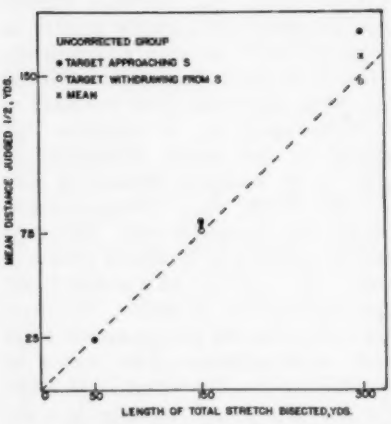


FIG. 2. The relationship between the point judged half-way between *O* and a marker, and the distance of the marker in yd. The line at 45° on the graph represents accurate fractionation.

Figure 3 shows the error that resulted from the fractionation of distance stretches as the distance of those stretches from *O* increased. The points connected by solid lines represent bisections of 50-yd. stretches. The error for each bisection is plotted at the true half-way point of the stretch bisected. For instance, the point at 125 yd. represents a bisection of a stretch extending from 100 yd. to 150 yd. The dotted lines connect error values that resulted from dividing a 75-yd. stretch into intervals subjectively equal to one-third and two-thirds.

The horizontal line drawn at zero error indicates the expected judgments if "constancy" for distance stretches were perfect. The dashed curve, on the other hand, describes the judgments expected from a hypothetical *O* (6 ft. tall) who fractionated the stretches by making the two fractions subtend (*a*) equal visual angles in the case of halving 50 yd., or (*b*) visual angles equal to one-third and two-thirds of the total angle subtended by 75 yd. The obtained judgments are close to the line of objective accuracy and do not follow the hypothetical visual angle judgment curve. The judgments at the far stretches show some tendency to be negative. Since the error is small and since the two hypothetical curves are close together at the far stretches, it is not possible to conclude that the plotted points follow either one or the other curve at 175, 200, and 250 yd. The points plotted at 25, 50, 100, and 125 yd. follow the "constancy" line.

This demonstration suggests the interdependence of all spatial constancies. All the dimensions of a surface, both frontal dimensions (termed size) and the third dimension (termed depth or distance) show constancy of magnitude. Gibson's more

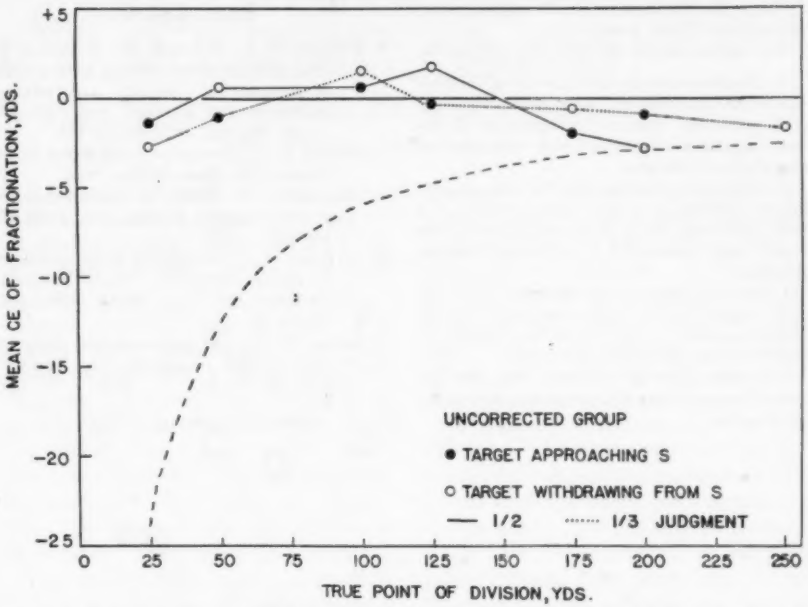


FIG. 3. The group mean CE of (*a*) halving a 50-yd. stretch (solid line) and (*b*) dividing a 75-yd. stretch into one-third and two-thirds (dotted line) as a function of the true division point in yd. The dashed curve represents the expected error if *O* were to divide the stretches so that the resulting fractions subtended appropriate fractions of visual angle.

general statement may thus be applicable—that "Scale, not size, is actually what remains constant in perception" (**2**, p. 181).

The effect of correction on fractionation.—So far, only the results of uncorrected *O*s have been considered. Another large group of *O*s whose errors were corrected as the judgments proceeded can be compared with them. Table 2 gives mean CE's and *SD*'s for both groups. The corrected group had 12 of 20 CE's significantly greater than zero, while only 9 of the uncorrected group's were significantly greater. A *t* test comparing the magnitude of the CE's of the corrected and uncorrected groups shows that five of the CE's are significantly larger ($p < .05$) in the corrected group (at 125 and 200 yd. with bicycle withdrawing, and at 25, 200, and 250 yd. with bicycle approaching). The mean CE is in no case significantly larger in the uncorrected group. Correction, therefore, does not appear to increase the accuracy of fractionation, which is, indeed, very good already. The fractionation with correction did, however, serve as an effective training procedure for later absolute estimates of distance (**1**).

SUMMARY

The present study is an application of a psychological scaling operation (fractionation) to perceived distance. A 300-yd. stretch of grass-covered flat ground was used for *O*'s judgments. A bicycle moved up or down this stretch, and *O* indicated his judgment by stopping the rider at what he considered the correct division point. Six bisections and four trisections were made, with different intervals and ground stretches so chosen that one of the true division points would

lie every 25 yd. from zero (*O*'s station point) through the 250-yd. point.

The results suggest the following conclusions:

1. Observers can divide stretches of distance (up to 300 yd.) into halves or thirds with very good accuracy. Perceived magnitudes of distance appear to correspond well with physical magnitudes of distance.

2. There is no tendency for the judgments to follow a hypothetical curve which would result from their being based on the magnitude of the visual angle subtended by different distance stretches.

3. Error is related to the direction of motion of the target; the CE tends to be positive as the bicycle approaches, and either less positive or negative when the bicycle withdraws.

4. Fractionation of distances was not improved by correcting *O*'s errors, nor was variability reduced.

References

1. Gibson, E. J., Bergman, R., & Purdy, J. The effect of prior training with a scale of distance on absolute and relative estimation of distance over ground. *J. exp. Psychol.*, 1955, **50**, 97–105.
2. Gibson, J. J. *Perception of the visual world.* Boston: Houghton Mifflin, 1950.
3. Gilinsky, A. S. Perceived size and distance in visual space. *Psychol. Rev.*, 1951, **58**, 460–482.
4. Gilinsky, A. S. Perception of size of objects at various distances. *USAF Personnel Training Res. Cent. Bull.*, 1954, No. TR-54-92.
5. Smith, W. M. A methodological study of size-distance perception, *J. Psychol.*, 1953, **35**, 143–153.

(Received January 1, 1955)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

EFFECTS OF COMPLEX TRANSFORMATIONS OF FEEDBACK UPON SIMPLE INSTRUMENTAL BEHAVIOR[1]

CLYDE E. NOBLE AND IRVIN G. BROUSSARD

Louisiana State University

Research in human instrumental behavior has shown that knowledge of results or psychological feedback (3) is a major factor in the establishment and maintenance of response strength as well as proficiency or skill. Following early qualitative work by Judd (5), Arps (1), and Thorndike (6), later investigators became interested in quantitative aspects of the cue, drive, and reward properties of the law of effect. In line drawing, for example, Trowbridge and Cason (7) demonstrated the dependence of skill upon specificity of feedback. When the direction and magnitude of error was reported, *Ss'* responses were more accurate and goal scores were more probable than when *E* merely said "right" or "wrong."

Recently, Bilodeau (2) studied the influence of transformations of feedback on the behavior of turning a micrometer to obtain a certain score. Linear equations of the form $Y = aX + b$ were employed to convert *Ss'* true scores (X) into reported scores (Y). The arbitrary goal or incentive was a reported score of 200, defined as a scale reading of .2 in. multiplied by a constant of 1000. By appropriate choice of slope and intercept parameters in his design, Bilodeau caused all lines to intersect at the point $Y = X = 200$. Each *S's* actual deviations from the target (X-200) were calculated and the mean performance plotted over 16 successive practice trials. The results indicated: (*a*) a family of stable error curves which approached a common asymptote as a limit, (*b*) a rate of approach to mastery which varied positively with the magnitude of the slope and negatively with the intercept, and (*c*) response variability which increased initially and then decreased as practice continued.

A technical problem discussed by Bilodeau is the fact that his conditions entailed joint, inverse variation of the slope and intercept parameters. If either *a* or *b* had been held constant while the other varied, a fixed incentive would have been logically impossible and the tasks psychologically incomparable. Bilodeau concluded, therefore, that proficiency is positively related to the magnitude of the reported target deviations (Y-200). The present experiment represents an extension of Bilodeau's study in that micrometer performance will be investigated under nonlinear conditions of feedback. This is intended to accomplish two primary objectives: (*a*) the resolution of the problem of multiple parameters, and (*b*) the systematic comparison of linear and curvilinear transformations.

With respect to the first objective, we have chosen a parabolic equation of the general form $Y = aX^b$ for all experimental conditions. The mathematical properties of this function are such that the slope parameter (*b*) may be varied while holding constant the origin (log *a*) and goal intercept (unity). Three treatments, differing quantitatively in the manner in which X scores are converted to Y scores, are sufficient for a repre-

[1] This paper is based on a master's thesis investigation by the junior author under the guidance of the senior author.

FIG. 1. Experimental design for studying the effects of curvilinear transformations of feedback.

sentative analysis. The transformation equations are as follows:

Group	Variable	Equation	
Minus	Negative acceleration	$Y = X^{.25}$	(1)
Zero	Constant velocity	$Y = X^{1.0}$	(2)
Plus	Positive acceleration	$Y = X^{4.0}$	(3)

Figure 1 shows the experimental design graphically. It will be observed that Group Zero is linear and equivalent to Y = X. This condition serves as a point of reference for evaluating the other two treatments. It represents the usual case of psychological feedback; i.e., true information. The reported scores of Group Minus are a negatively accelerated increasing function of their actual scores, whereas the reported scores of Group Plus are a positively accelerated increasing function of their actual scores. Note that the slopes of all three equations increase in the order shown and that the b parameters of Equations 1 and 3 are simple reciprocals of each other.

With respect to the second objective, the present study will permit comparisons between ordinary feedback and two symmetrical complex transformations. Equation 2 provides both a control group and a link with the Y = X condition of the Bilodeau experiment, whereas Equation 1 should elicit behavior similar to that of his Y = .500X + 100 group.[2] There is no linear parallel in Bilodeau's study for our Equation 3, essentially because he did not employ a negative Y intercept. (An analogous equation might have been Y = 2.0X − 200, but this would have entailed large, negative reported scores for the early trials.) Considering the differences in initial reported scores and subsequent feedback rates of the three groups in Fig. 1, it may be expected that Group Plus will at first overshoot the target and then improve in accuracy by discrete corrections. Generalizing from Bilodeau's results, Groups Zero and Minus should achieve proficiency more smoothly, with Group Zero definitely superior.

METHOD

Subjects.—The *S*s consisted of 96 male and female undergraduate students at Louisiana State University, assigned independently and without bias to the three treatment groups ($n = 32$). The male/female ratio for each condition was as follows: Minus = 22/10, Zero = 21/11, Plus = 18/14. None had previous experience with the apparatus and all were tested singly. The *S*s' ages ranged from 18 to 31 yr. with a mean of 20.7 yr. There were no eliminations.

Apparatus and procedure.—A 3–4-in. Sears Craftsman micrometer was mounted in a plywood screen in the manner described by Bilodeau (2). The *S*'s problem was to learn to turn the exposed thimble by an amount necessary to receive a numerical score of 200. The accuracy of *S*'s response on a given trial, therefore, was contingent upon *E*'s report of his performance on the preceding trials. The instructions explained that 200 constituted the goal score, that size of score increased with the number of turns, that each trial would start from zero, and that a uniform type of response was required throughout practice. To remove conflicting visual cues and to simplify *E*'s recording task, our *S*s were blindfolded with occlusion goggles and were required to use the left hand to manipulate the thimble. (In Bilodeau's procedure *S*s were not blindfolded and used the right hand.) The vernier scale was read by *E* after each response or series of responses (trial), then reset to zero for the succeeding trial.

[2] In order to facilitate comparisons with Bilodeau's design, Fig. 1 has been constructed on the basis of a target score of 200 rather than unity. This was achieved by the following simple modifications of Equations 1 and 3 in order to refer all curves to a common scale:

$$Y = 200\,(X/200)^{.25} \quad (1a)$$

$$Y = 200\,(X/200)^{4.0} \quad (3a)$$

Approximately 5 sec. elapsed between *S*'s response and *E*'s report, followed by 10 sec. before the next trial. All *S*s were given 20 practice trials to achieve the incentive.

Another procedural departure from the Bilodeau study was the restriction of *S*'s first response to a *half* turn, rather than a *full* turn. A response of exactly 180° equals a true score of 12.5 (in units of .001 in.). Thus, for all treatments 16 half turns were necessary and sufficient to achieve a goal score of 200. Of course, *S*s were not informed of this fact and care was taken to isolate them from the remaining pool of *S*s upon completion of the experiment. As can be seen in Fig. 1, this method of standardizing Trial 1 was required in order to provide ample opportunity for differential acquisition rates. Given an actual score of 16, the reported scores for the various groups would be as follows: Minus = 113, Zero = 20, Plus = .02. The so-called "practice trial" also served effectively to match *S*s as to initial means and variances, since a single wrist motion constitutes a well-standardized response unit. The *E* recorded the X scores in terms of .001 in., then rounded to the nearest 10 before entering the appropriate transformation table. By this device all groups had 20 available Y scores between 0 and 200.

Results and Discussion

Acquisition phenomena.—In order to observe the directionality of errors, it is necessary to plot the data in terms of algebraic deviations (X-200) from the goal score. Bilodeau calculated mean absolute deviations, but this was because his *S*s made few responses greater than 200. Furthermore, since skewness was an expected characteristic of the frequency distributions, especially those of Group Plus, medians were computed instead of means.

Figure 2 presents the error curves as a function of successive trials for the three transformation groups. The initial X scores are almost exactly equal. The means and *SD*'s (in parentheses) for the respective groups on Trial 1 are as follows: Minus = 17 (3.0), Zero = 16 (3.8), Plus = 16 (3.5). A simple analysis of variance

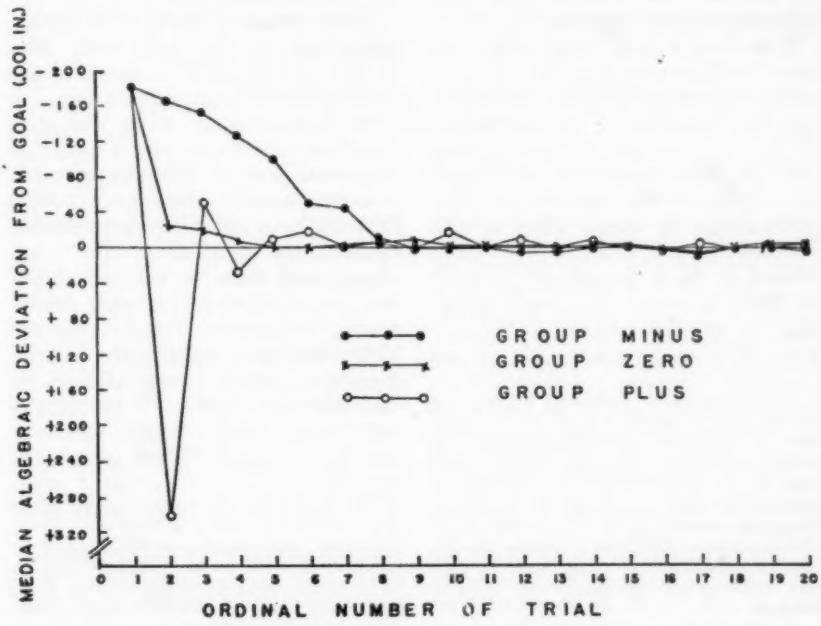


Fig. 2. Median algebraic deviations of true scores from the goal for the three experimental groups as functions of practice.

for 2 and 93 *df* was performed upon these data, resulting in an *F* ratio less than unity. The over-all mean of 16.33 represents an overestimation of one half turn by 30.6%.

Inspection of the performance characteristics shows that the error curve for Group Minus decreases at a much slower rate than that of Group Zero, which is negatively accelerated throughout. As Bilodeau found for his comparable groups, "transformation" by Equation 2 is superior to that by Equation 1. However, the rate of acquisition for the Y = X condition in the present experiment is much faster than in Bilodeau's. Among possible differentiating factors, the most likely one is that our *S*s were required to make a 180° turn on Trial 1, whereas Bilodeau's *S*s were instructed initially to turn the thimble 360°. This phenomenon may be attributed to the motivational effect of *S*'s reported distance from the goal following his first response.[3]

From Fig. 2 it is evident that the curve for Group Minus has an early positive acceleration followed by later negative acceleration. On the whole, the curve has a definite sigmoid form and closely resembles Bilodeau's Y = .500X + 100 condition. The performance of Group Plus is consistent with expectation in that Trial 2 reveals a large overshooting error (+300.5) followed by a small undershooting error (− 52.8) on Trial 3.[4] There is a subsequent trend toward diminishing oscillation, but the alternation is irregular and slightly asymmetrical in favor of undershooting.

The accuracy of Group Zero is superior throughout. From Trial 4 onward, the mean median deviation scores for the respective groups are: Zero = − 1.45, Plus = − 2.16, Minus = −18.55. Descriptively and by analogy to military gunnery methods, the behavior of Group Zero can be called a "laying-on" technique, while that of Group Minus is characterized by "creeping" toward the target. Distinctively different from either of these is the "bracketing" procedure of Group Plus, whereby successive observations of "strikes" permit the eventual "zeroing-in" of the target. More generally, the performances of Groups Zero and Minus may be termed a continuous approach, whereas that of Group Plus represents a discontinuous approach.

Performance variability.—A further comparison of our data with Bilodeau's can be made in terms of the *SD*'s of the X scores. In Fig. 3, *SD*'s of the true scores are plotted as functions of trials for the three transformation groups. The general trend is an increase in variability on Trial 2 followed by a fairly uniform decrease to the limit of training. The magnitude and form of the variability function is different for each experimental treatment. Group Plus exhibits maximum variability on Trial 2 which is almost a hundred times its variability on Trial 1. Although the *SD*'s drop rapidly in the following four trials, Group Plus is more vari-

[3] Other procedural differences between the two experiments (such as right vs. left hand, vision vs. nonvision, and military personnel vs. college students) are probably of little importance in the light of available evidence. A preliminary micrometer study in the L.S.U. laboratory employing the Y = .500X + 100 transformation, a 360° Trial 1 response, the left hand, blindfolding, and college level *S*s gave data which agreed very closely with those of Bilodeau.

[4] A problem arose in this connection due to physical limitations of the apparatus. Since a 3–4-in. micrometer permits only a 1-in. excursion of the spindle from the zero setting, the maximum possible X score is 1000 (1 in.). About 31% of the *S*s in Group Plus on Trial 2 turned the spindle to the limit, giving rise to indeterminate mean scores. Exactly 12.5% of Trial 3 responses and about 3% of Trial 4 responses were of this class. After Trial 4 the problem did not exist, but it was necessary in the foregoing cases to record an arbitrary X score of 1000.

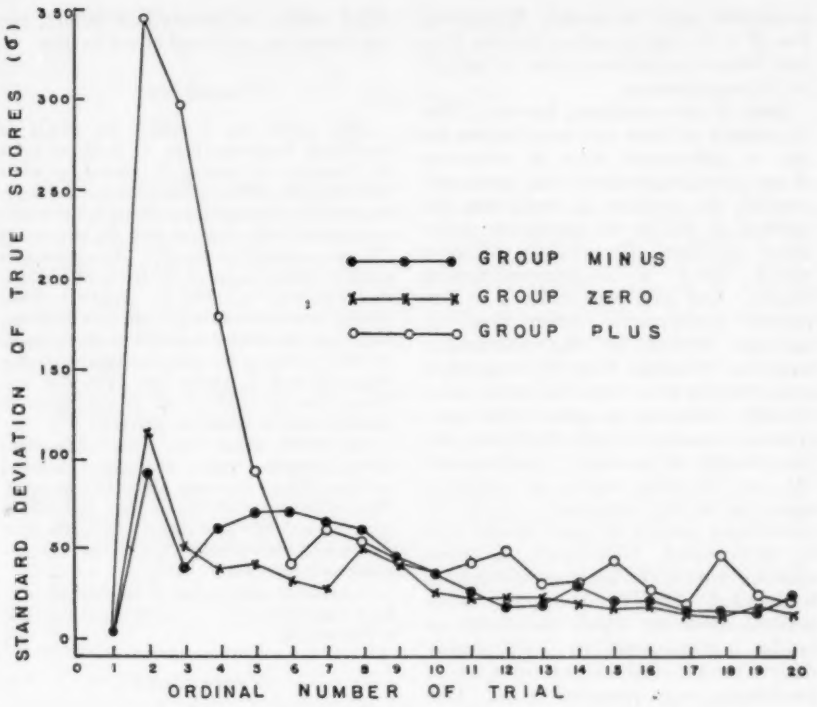


FIG. 3. Standard deviations of true scores for the three experimental groups as functions of practice.

able than the others for the last half of the experiment. Groups Zero and Minus are comparable initially, their Trial 2 *SD*'s being about 30 times that of Trial 1. Group Zero's variability decreases with practice, but Group Minus undergoes a rise at Trial 4 which peaks two trials later and then gradually falls. All curves appear to be merging toward the same asymptote. It is notable that the frequency distributions become more symmetrical as the terminal trials are approached.

In his analysis of the effects of linear score transformations upon micrometer performance, Bilodeau hypothesized that *S*s respond initially as if they expected the functional relationship between reported and true scores to be Y = X, and that they subsequently extinguish this pre-experimental assumption "at a rate which varies with the magnitude of the transformation constants" (2, p. 420). To evaluate this, he tested the mean differences between each *S*'s aspired and achieved X scores on Trial 2 for significance. Of the five independent comparisons, only two t ratios were small enough to permit retention of a null hypothesis ($p > .05$). Bilodeau's contention was not even supported for his Y = X group; with a $p < .01$, *S*s receiving no transformation deviated maximally from prediction. We applied the same test to our data, calculating individual Trial 2 expectancies from the response and feedback on Trial 1. The mean difference for Group Plus was indeterminate due to apparatus restrictions mentioned above, but the obtained scores of Groups Zero and Minus were entirely

consistent with Bilodeau's hypothesis. For $df = 31$, the p values for the Zero and Minus comparisons were $> .60$ and $> .10$, respectively.

Even if one concludes, however, that the results of these two experiments are due to differential rates of extinction of an initial equivalence set, there still remains the problem of explaining the manner in which the extinction mechanism operates. Presumably, Bilodeau would favor a counterconditioning theory. Our own preference is to emphasize motivational rather than associative factors in the micrometer situation. Assume that the magnitude and direction of the reported target error (Y-200) following a given trial constitutes an aversive state of affairs. By the principle of secondary reinforcement (4), any X score causing an apparent reduction of this discrepancy and better orientation toward the goal should tend to be repeated. Conversely, X scores that are wide of the mark should tend to be avoided. For this reason, the mathematical form by which knowledge of results is transmitted to *S* will largely determine his average rate of gain in proficiency over repeated trials. This interpretation of the role of feedback transformations regards the learning factor as a constant with incentive-motivation as the variable. This is consistent with special effects like the relatively slow acquisition of Bilodeau's $Y = X$ group, the initial overshooting and high variability of our $Y = X^{4.0}$ group, and the superiority of $Y = X^{1.0}$ over $Y = X^{.25}$.

The differences found in the present investigation can be attributed to a single experimental factor, but it would be erroneous to conclude that proficiency is a direct function of the size of the b parameter. Differential acceleration components in psychological feedback affect human instrumental behavior in a complex manner. Our findings suggest that the informative or cue properties of knowledge of results should not be stressed at the expense of its drive and reward properties. The micrometer offers excellent opportunities for studying these problems, in addition to such allied topics as success and failure, ego involvement, and level of aspiration.

Summary

This paper has described the effects of curvilinear transformations of feedback upon the behavior of turning a micrometer while blindfolded to achieve a certain numerical score. Parabolic equations of the general form $Y = aX^b$ were employed to convert each *S*'s true scores (X) into reported scores (Y). This procedure entailed manipulation of the factor of acceleration (positive vs. zero vs. negative), while holding constant both origin and goal intercept. Ninety-six *S*s, divided equally into three experimental conditions, were differentiated as Group Minus ($Y = X^{.25}$), Group Zero ($Y = X^{1.0}$) and Group Plus ($Y = X^{4.0}$). All *S*s were given 20 practice trials to attain the goal score.

In general, Group Zero (true information) showed superior accuracy throughout, followed by Group Plus and Group Minus in that order. The performances of Groups Zero and Minus were characterized by a continuous approach to the target, while Group Plus exhibited a discontinuous approach.

A tentative explanation of the role of feedback transformations in instrumental behavior was presented.

References

1. Arps, G. F. A preliminary report on work with knowledge versus work without knowledge of results. *Psychol. Rev.*, 1917, **24**, 449–455.
2. Bilodeau, E. A. Speed of acquiring a simple motor response as a function of the systematic transformation of knowledge of results. *Amer. J. Psychol.*, 1953, **66**, 409–420.
3. Brown, J. S. A proposed program of research on psychological feedback (knowledge of results) in the performance of psychomotor tasks. *USAF, Hum. Resour. Res. Cent., Conf. Rep.*, 1949, No. 49–2.
4. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
5. Judd, C. H. Practice without knowledge of results. *Psychol. Rev. Monogr. Suppl.*, 1905, **7**, 185–198.
6. Thorndike, E. L. The law of effect. *Amer. J. Psychol.*, 1927, **39**, 212–222.
7. Trowbridge, M. H., & Cason, H. An experimental study of Thorndike's theory of learning. *J. gen. Psychol.*, 1932, **7**, 245–258.

(Received January 6, 1955)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

DISCRIMINATION DECREMENT AS A FUNCTION OF TIME IN A PROLONGED VIGIL[1]

PAUL BAKAN[2]

New York University

This experiment was designed to investigate the change in differential brightness threshold as a function of time, in a situation where *S* is required to respond overtly to irregularly presented variable stimuli differing in brightness from a regularly recurring standard stimulus.

Mackworth (**3, 4**) found a decrease in the number of successful responses as a function of time spent at the Clock Test, a task requiring continuous vigilance. In the Clock Test, *S* is required to respond overtly to irregularly and infrequently presented critical stimuli, i.e., double-length movements of a pointer in front of a flat white surface. The pointer jumps regularly once a second, making a single-length movement, with an occasional double-length movement occurring in place of a single-length movement. The *S* is not required to make any overt response to single-length stimuli. The decrease with time in the number of responses to critical stimuli in such a situation may be interpreted in terms of an increase in threshold for the required discrimination in the course of time. In order for *S* to respond overtly to the critical double-length stimulus, he must discriminate between the single-length stimulus and the double-length stimulus. If the decrement in frequency of response is thought of in terms of an increase in threshold, then it should be possible to obtain a measure of decrement by taking repeated threshold measures during a prolonged discrimination task. In this experiment the method of limits is used to measure thresholds over a period of time in a continuous brightness discrimination task, to see whether a decrement can be measured in terms of an increase in threshold.

METHOD

Subjects.—Twenty paid *S*s (10 male and 10 female), ranging in age from 16 to 35 years, participated in this experiment. Each *S* was tested on two different days, generally during the same week.

Test situation.—The *S* viewed a test area, 15 min. of arc in diameter and 9 ft. away, located within a uniformly illuminated surround (2.6 ft.-L.). The test area flashed regularly once a second, holding at a base brightness (2.1 ft.-l.) for ½ sec. and flashing to a standard brightness (3.1 ft.-L.) for the remaining ½ sec. During a trial the test area flashed to a level brighter than standard and *S* was instructed to push a button whenever he detected a stimulus brighter than standard.

In order to study the course of the threshold change during the vigil, the brightness increment

[1] This paper represents part of a dissertation submitted in 1953 to the Graduate School of New York University in partial fulfillment of the requirements for the Ph.D. degree. The writer wishes to acknowledge his indebtedness to Lawrence Karlin of New York University, who supervised the thesis.

The research was done at the Training Research Laboratory of the University of Illinois. It was supported in part by the United States Air Force under contract No. AF 33(038)-25726, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government. A report of the contents of this paper was given at the 1953 meeting of the Midwestern Psychological Association.

The writer wishes to express appreciation to William E. Kappauf and Lyle H. Lanier of the University of Illinois for their invaluable aid during the course of this study.

[2] Now at Michigan State University.

required to evoke a response was determined by the ascending method of limits on each trial. When a trial was called for, a test stimulus of Brightness Level 1, the smallest brightness increment over the standard, was presented in place of the standard. If no response occurred, Brightness Level 2 was presented 5 sec. later, i.e., five flashes later, and so on with increasingly brighter test stimuli at 5-sec. intervals until *S* indicated seeing a brighter stimulus by pressing the button.

The test stimuli ranged from 3.2 to 4.0 ft.-L. in a series of nine equal log steps. Trials were begun by *E* at times which had been scheduled in advance. Test sessions were 1.5 hr. long and the distribution of trials in time was such that there were six equivalent 15-min. periods. Six trials were scheduled for each 15-min. period. The trials were irregularly spaced and came after waits of 50, 60, 65, 140, 185, and 350 sec. between the beginning of one trial and the beginning of the next. Four different 15-min. programs of waits were used. These programs were chosen at random from the possible arrangements of the six wait periods. The varied programs were used so that no two successive 15-min. periods would have the same wait sequence, and so that the wait sequences on the second day would be different from those on the first day for any individual *S*.

Practice.—After the instructions to *S*, various test stimuli ranging down to *S*'s threshold were demonstrated to *S*. This was followed by a 7.5-min. practice period, during which five trials were given under test conditions like those of the experimental session. A 5-min. rest period followed this practice period.

Scoring.—Two types of score were obtained from the data collected during the one and one-half hour vigil: (*a*) a measure of the mean brightness increment required to evoke a response for each 15-min. period, i.e., the mean difference limen,[3] and (*b*) a tally of errors of commission, for each 15-min. period, i.e., a record of responses made when no brighter stimulus was presented.

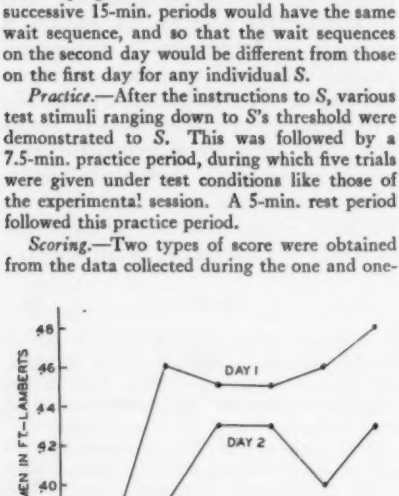


FIG. 1. The effect of time on the difference limen for each day.

TABLE 1

ANALYSIS OF VARIANCE FOR THRESHOLD DATA

Source	*df*	Mean Square	*F*	*P*
Periods (P)	5	4.10	8.72	.001
Days (D)	1	13.54	7.78	.05
Ss	19	7.39	4.25	.01
P × D	5	.43	—	
P × *Ss*	95	.39	—	
D × *Ss*	19	1.74	3.70	.001
P × D × *Ss*	95	.47	—	

RESULTS

Change in threshold.—The change with time in the difference limen is shown in Fig. 1 for each day. The threshold increase was such that in order for a brighter stimulus to be detected, it had to be about 25% more intense late in the session than at the beginning. On the basis of the analysis summarized in Table 1, it is clear that the variance contributed by 15-min. periods is highly significant.

Difference between days.—The variance contributed by days, when tested against the significant day-by-subject interaction, is significant at the 5% level. As Fig. 1 shows, performance during each 15-min. period was better on Day 2 than on Day 1. Clearly there was a practice effect which improved the general level of performance on Day 2, but this effect does not seem to have significantly altered the decremental trend which is found for both days.

[3] In four of the 1,440 trials given, no response was evoked for the largest brightness increment (Level 9). These trials were scored as if a response had been evoked at Level 9.

Individual differences.—The variance contributed by individual *S*s was significant at the 1% level, indicating that in terms of over-all threshold scores some *S*s were good at the task and others poor. Further, the *S*-by-day interaction was significant, indicating individual differences in the extent of change in performance from Day 1 to Day 2.

An analysis of the difference between "good" and "poor" *S*s was carried out. Good *S*s may differ from poor *S*s in either or both of two ways: they may differ because of differences in threshold for the required discrimination, or they may differ because of differences in the course of threshold change with time. In order to determine whether differences in the course of threshold change with time are related to differences in over-all ability, *S*s were ranked according to ability on the basis of the mean threshold over all 15-min. periods on both days, and divided into two equal groups, good and poor. The effect of time on threshold is shown for each group in Fig. 2. The ability-by-period interaction was significant at the 5% level. This means that associated with the difference in over-all ability is a difference in the course of threshold change with time. As can be seen from Fig. 2 the good *S*s show, on the average, a partial recovery from decrement after about 45 min. at the task (difference between Period 3 and Period 5 significant at 5% level), whereas the poor *S*s show, on the average, a continuing increase in threshold in the course of the test session. Mackworth (**4**) also found that the slope of decline of performance with time was less marked with good *S*s than with poor *S*s.

Errors of commission.—For each day there was a trend toward a decreasing number of commission errors from the beginning to the end of the session. Also, there were fewer commission errors on Day 2 than on Day 1; the difference between days was significant at the 5% level. On Day 1 the mean number of commission errors decreased from 13 in the first period to 6 in the last period. On Day 2 there was a decrease from a mean of 8 in the first period to 3 in the last period. Although 15 of the 20 *S*s made fewer commission errors on Day 2, the extent of the day-to-day change was markedly influenced by four *S*s with large differences in commission error frequencies between days.

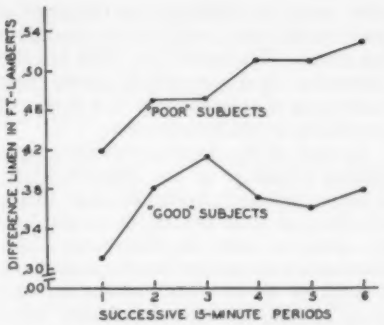


FIG. 2. Difference limen as a function of time for "good" and "poor" *S*s classified on basis of over-all ability. First and second days have been combined.

DISCUSSION

The results show that a decrement can be observed by taking repeated threshold measures during a prolonged discrimination task. Thinking of the decrement as a threshold change not only provides a setting for all decrement data but also leads to the assertion, supported in the literature, that if the discrimination is made sufficiently easy, then there should be little or no decrement in frequency of response in the course of time.

A possible explanation of the increase in threshold with time is suggested by some observations of Pavlov (**5**, p. 307)

who noted a tendency on the part of dogs to fall asleep while in the conditioning stand. Kleitman (2, p. 286) has also presented evidence which shows that monotony of surroundings is a means of producing drowsiness or sleep.

In view of the monotony of the experimental situation of the present study, it seems plausible to assume that this is the type of situation which would tend to induce a state of drowsiness. This assumption seems justified in view of the reports of *S*s in which monotony and sleep symptoms were frequently mentioned. Of interest in this connection is the fact that it was possible to eliminate decrement in the Clock Test by having *S*s swallow a Benzedrine tablet before the test session (3, 4). In the case of motivated human *S*s, set to continue working at the task, the tendency to drowsiness and sleep may be fought by an attempt on the part of *S* to change the situation psychologically in order to to reduce the monotony. This may be done by means which reduce attentiveness to the discrimination task (e.g., daydreaming) and thus impair discrimination performance. Reports from *S*s on their activities during the test session support this explanation. Broadbent (1) has recently formulated a similar analysis of watchkeeping tasks in terms of attention.

Summary

This study utilizes a method of repeated threshold measurement during a prolonged vigil in order to measure vigilance decrement.

Twenty *S*s were tested individually for 1.5 hr. on each of two days. The *S* viewed a test area which flashed regularly once per second, to a standard (3.1 ft.-L.) brightness level for ⅛ sec. Test stimuli brighter than standard could be substituted for the standard by *E* at predetermined times. Test stimuli ranged from 3.2 ft.-L. to 4.0 ft.-L. in a series of nine equal log steps. The *S* responded by pressing a key whenever he saw a brighter flash. Test stimuli were used to measure threshold by the method of limits. During a trial, test stimuli were presented at 5-sec. intervals in an ascending series until an overt response was evoked.

In the course of the test session the difference between the standard and the variable had to be increased in order for responses to be elicited. The extent of this increase represents a measure of vigilance decrement. This decrement is interpreted in terms of response to irrelevant stimuli as a means of reducing monotony in a monotonous situation.

References

1. Broadbent, D. E. Classical conditioning and human watch-keeping. *Psychol. Rev.*, 1953, **60**, 331–339.
2. Kleitman, N. *Sleep and wakefulness.* Chicago: Univ. of Chicago Press, 1939.
3. Mackworth, N. H. The breakdown of vigilance during prolonged visual search. *Quart. J. exp. Psychol.*, 1948, **1**, 6–21.
4. Mackworth, N. H. *Researches on the measurement of human performance.* London: H. M. Stationery Office, 1950. (Med. Res. Council, Special Rep. Ser. No. 268.)
5. Pavlov, I. *Lectures on conditioned reflexes.* New York: International, 1928.

(Received January 12, 1955)

Journal of Experimental Psychology
Vol. 50, No. 6, 1955

THE NECESSARY CONDITIONS FOR CUE–POSITION PATTERNING

DAVID BIRCH AND VICTORIA VANDENBERG
University of Michigan

Theory building in discrimination learning has a history of conflicts and confusions concerning the most fruitful way to define the effective stimuli in a discrimination learning problem (**4, 6, 7, 9, 12**). The present study adopts the analysis of Spence, who describes the total stimulus situation in terms of absolute and patterned stimulus components, and attempts to clarify the conditions necessary for the occurrence of cue-position patterning (e.g., Black–Left) in a discrimination learning problem. In his discussion of the patterning of stimulus components Spence states, "Discrimination will involve such patterning according to this conception only when no one of the cue members is systematically reinforced more than the others" (**10**, p. 90). While the conditions stated by Spence may be very conducive for the occurrence of patterning, it may be that they are not necessary.

The design of this investigation presented the animal with a discrimination problem which could be learned on the basis of absolute components alone, but also allowed a test of the possible occurrence of stimulus patterning. The finding of stimulus patterning in this situation would indicate that Spence's conditions are not necessary ones.

METHOD

All *S*s were first required to learn a discrimination problem which paired Black on the Left with Gray on the Right ($\overset{+}{B}$ $\overset{-}{G}$) and White on the Right with Gray on the Left ($\overset{-}{G}$ $\overset{+}{W}$). In Table 1 this problem is shown as ($\overset{+}{B}$ $\overset{-}{G}$) and ($\overset{-}{G}$ $\overset{+}{W}$) designated Task 1. An analysis of the absolute stimulus components in the fashion of Spence shows S_B+, S_W+, S_G-, $S_R\pm$, and $S_L\pm$, and solution of the problem occurs when the combination of the approaching tendencies to S_B and S_L exceeds the combination to S_G and S_R, and when the combination to S_W and S_R exceeds that to S_G and S_L. Despite the fact that combinations of approaching and nonapproaching tendencies to these absolute components in the situation may account for the learning of the discrimination, it is possible that approaching tendencies to the patterned stimuli S_{BL} and S_{WR} and nonapproaching tendencies to S_{GL} and S_{GR} could have developed.

A series of transfer tasks using four groups of *S*s were run to test for the presence of differential approaching tendencies to the absolute and the patterned stimuli. Table 1 summarizes the problems presented to *S*s and provides analyses of the stimulus components for the learning tasks with and without patterning.

The performance of Group A on Task 2 was contrasted with that of Group B on Task 3. If learning to the *patterned* stimuli has occurred in the original training, Group A is expected to demonstrate significantly more correct choices than Group B on the initial segment of the transfer problem trials; if learning has been to absolutes alone, both groups should respond with approximately 50% correct. Learning to the *absolute* components during the original training should be manifested by significant differences in the initial number of correct responses between Group C on Task 4 and Group D on Task 5, with Group C showing the much larger number of correct choices.

Groups C and D were employed in a second test for learning with respect to patterned stimuli by training the two groups to criterion on Tasks 4 and 5 and then testing them on Task 6, which is the same as Task 2 for Group A. Group C should respond in the neighborhood of 50% correct on Task 6 whether or not the learning in Tasks 1 and 4 involved patterning, whereas Group D should respond approximately 50% correct if absolutes alone are involved but should be close to 100% correct choice if approaching tendencies to S_{BL} and S_{WR} on Task 1 and nonapproaching tendencies to S_{BR} and S_{WL} on Task 5 were acquired.

Subjects.—The *S*s were 28 experimentally naive, male rats, 90 to 120 days old at the be-

TABLE 1

SERIES OF LEARNING AND TEST TASKS WITH ANALYSES OF THE STIMULUS COMPONENTS WITH AND WITHOUT PATTERNING FOR THE LEARNING TASKS

Task	Groups	Analysis by Absolutes Alone	Analysis Including Cue-Position Patterning
1. $\binom{+}{B}\ \binom{-}{G}$ / $\binom{-}{G}\ \binom{+}{W}$ Learn	A, B, C, D	S_B+, S_W+, S_G- $S_R\pm$, $S_L\pm$	S_B+, S_W+, S_G- $S_R\pm$, $S_L\pm$ $S_{BL}+$, $S_{WR}+$, $S_{GL}-$, $S_{GR}-$
2. $\binom{+}{B}\ \binom{-}{B}$ / $\binom{-}{W}\ \binom{+}{W}$ Test	A		
3. $\binom{-}{B}\ \binom{+}{B}$ / $\binom{+}{W}\ \binom{-}{W}$ Test	B		
4. $\binom{-}{G}\ \binom{+}{B}$ / $\binom{+}{W}\ \binom{-}{G}$ Learn	C	S_B+, S_W+, S_G- $S_R\pm$, $S_L\pm$	S_B+, S_W+, S_G- $S_R\pm$, $S_L\pm$ $S_{BR}+$, $S_{WL}+$, $S_{GL}-$, $S_{GR}-$
5. $\binom{+}{G}\ \binom{-}{B}$ / $\binom{-}{W}\ \binom{+}{G}$ Learn	D	S_B-, S_W-, S_G+ $S_R\pm$, $S_L\pm$	S_B-, S_W-, S_G+ $S_R\pm$, $S_L\pm$ $S_{BR}-$, $S_{WL}-$, $S_{GL}+$, $S_{GR}+$
6. $\binom{+}{B}\ \binom{-}{B}$ / $\binom{-}{W}\ \binom{+}{W}$ Test	C, D		

ginning of the experiment. Fourteen animals were of a pigmented strain derived originally from the Lashley stock and maintained by the Psychology Department at the University of Michigan, and 14 were hooded rats of the Long-Evans strain obtained from the Lurie supply in Chicago. Groups A and C each contained four *S*s from the departmental stock and three from the Long-Evans strain; Groups B and D each contained three *S*s from the departmental stock and four from the Long-Evans strain. No difference in performance attributable to strain was observed. Four additional *S*s, two from each strain, were discarded because of stereotyped position habits on the original or transfer tasks.

Apparatus.—The apparatus was an open elevated **Y** maze, 2 ft. high, placed upon a table 4 × 5 ft. The table was 2½ ft. from the floor and located in the middle of a room 9½ ft. × 9½ ft. The experiment was conducted at night and illumination was provided by diffuse lighting from fluorescent lamps recessed in the ceiling. The runway was 32 in. long and 2 in. wide and led immediately to an enlarged choice platform about 8 in. in diameter. Removable stimulus stands of the same height as the runway and choice platform formed the arms of the **Y**. These stands were 10 in. long and 4 in. wide and were separated from the choice platform by a 5-in. gap. The distance between the stands at the narrow portion of the arms was 7½ in. and 19 in. at the wide portion. The entire stimulus stand was painted either black, white, or gray, and throughout the experiment two stands of each brightness were used interchangeably. The runway and choice platform were painted the same gray as the gray stimulus stands. A small removable metal food cup could be at-

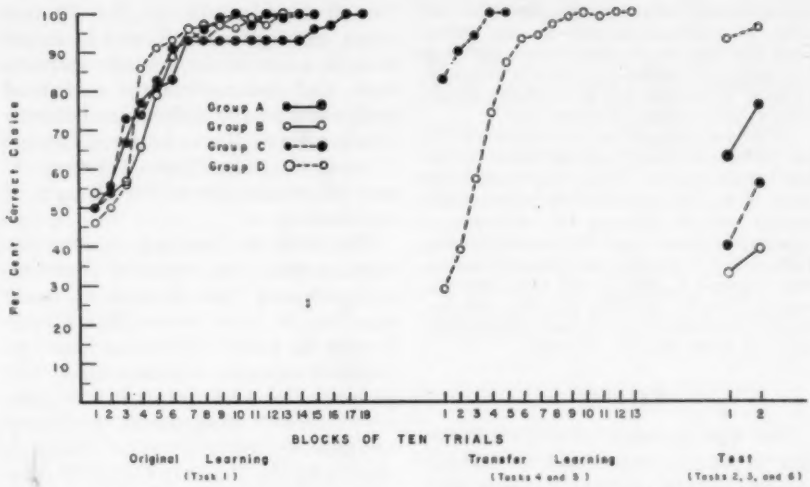


FIG. 1. Percentage of correct choices during the original learning, the transfer learning, and the test tasks.

tached at the far end about 1 in. below the level of each stimulus stand.

Adaptation training.—The *S*s were placed on a 24-hr. feeding schedule for about 2 weeks prior to the experiment. During this time they were handled extensively and trained to walk along a narrow board placed between two chairs, and to cross a gap up to 5 in. from one chair to the other to obtain food. The food consisted of Miller's Homogenized Dog Food pellets and the daily ration for all *S*s varied from 8 to 12 gm. depending on the size and need of the particular animal. The *E* attempted to estimate such needs during the adaptation period and felt this variation to be more effective in controlling excessive loss of weight and overactivity than a standard amount for all *S*s. Once determined, the amount of food given to any *S* was kept constant throughout the experiment.

Pretraining on the apparatus.—After the adaptation period all *S*s were given at least two days' experience, 5 trials each day, on the apparatus. The two stimulus situations in the original learning task, B-G and G-W, were presented an equal number of times and in random order. During this period, food cups with one pellet of food (approximately .07 gm.) in each cup were attached to both stimulus stands. Most *S*s ran down the runway, approached one of the stands, ate the pellet, crossed back to the choice platform and then to the other stand and ate the pellet there. No guidance was given to *S*s who approached both stimulus stands readily. In this way *E* could determine position preferences from the number of times *S* selected either the right or the left as his first choice. The *S*s who made but one choice were guided to the other side. Those who were more tense and hesitated in making any choice were guided alternately to both sides. These *S*s were given a few more pretraining trials on the apparatus until they reacted without any hesitation to both stimulus stands. This behavior was used by *E* as a criterion of "readiness." Analysis of the data showed that those *S*s ($N = 8$) who received more than 10 trials of pretraining on the apparatus did not differ from the others.

Discrimination training.—On the day following pretraining, *S*s were presented with the original learning task (Task 1). Ten trials per day, 5 to each of the stimulus pairs, were given by a series of Gellerman orders (3). If *S* approached the positive stand, he was allowed to eat one small pellet of food placed in the food cup. Immediately thereafter he was picked up and placed on a chair in the experimental room. If he approached the negative stand, he was picked up after he found neither food cup nor food at the end of the platform. No punishment was used, and no correction was permitted.[1] Usually four *S*s were run at a time, all *S*s com-

[1] One *S* in Group A who developed a position habit in the original learning task was allowed to explore both stimulus stands in both stimulus situations on 5 trials of the 15th day; he reached criterion within the following three days.

pleting one trial before the beginning of the next trial. If conditions necessitated the use of fewer than four *S*s, the time between trials was kept comparable to that when four *S*s were used. Criterion of learning was 20 successive correct trials on two successive errorless days.

At the beginning of the discrimination training, *S*s were randomly assigned to one of the four transfer groups. Those who reached criterion in the original task were presented the transfer task the following day. Criterion of learning for Groups C and D in transfer learning Tasks 4 and 5 was also two successive perfect days. Groups C and D were then placed in transfer Task 6. All tasks were run with one alternative rewarded and the other nonrewarded according to the schedule of Table 1.

Results

The performance of each of the four groups in percentage of correct choices for the original learning and each of the transfer tasks is shown in Fig. 1. Each point is based upon 10 trials, 5 involving black and 5 involving white, averaged over seven *S*s. The plots indicate no appreciable differences between the groups on the original learning but sizable differences between Groups C and D on the transfer learning and between all groups on the test trials. The slightly extended learning period for Group A on Task 1 is due to the position habit responses of one animal over Days 7–14.

Table 2 presents the mean number of errors per *S* per block of 10 trials for all four groups on the 20 test trials, and for Groups C and D on the first 20 trials of the transfer learning task, and summarizes the statistical evaluation of the differences between Groups A and B and between Groups C and D. The Mann-Whitney *U* test (**8**) was employed for all tests of significance.

The test for learning to absolute components, the transfer learning task, showed that Group C made significantly fewer errors than Group D with the greater difference found on the G-B stimulus combination. The 20 test trials used to test for learning to patterned components produced significantly fewer errors in Group A than in Group B and in Group D than in Group C.

In fact, except for the Group B-Group C difference, which reaches only the .15 level of significance, analysis of the total error score on the two test trial days shows all differences between pairs of groups to be significant beyond the .05 level. Thus, the ranking of the groups from least to most errors on the test trials may be taken with some confidence to be DACB. Examination of Table 2 reveals that in all but one comparison (A-D), the groups differed more on the W-W stimulus combination than on the B-B, but in every

TABLE 2

Mean Number of Errors per *S* per Block of 10 Trials on the 20 Test Trials and the First 20 Trials of the Transfer Learning Task Along with Significance Levels* for Differences Between Groups

Test Trials				Test Trials				Transfer Learning		
Group	B-B	W-W	Total	Group	B-B	W-W	Total	G-B	W-G	Total
A	3.4	2.7	3.1	C	4.6	5.9	5.2	0.1	2.6	1.4
B	4.4	8.4	6.4	D	0.1	1.0	0.6	8.9	4.4	6.6
p	0.58	0.002	0.002	*p*	0.006	0.002	0.000	0.000	0.42	0.002

* The values for the significance levels were obtained from Mann-Whitney *U* tests which were run on the number of errors during the 20 test trials and the first 20 trials for the transfer learning task. All probabilities are for two-tailed tests.

comparison except B-C the direction of the difference between groups was the same in both stimulus combinations.

Differences between pairs of the four groups in the total number of errors made in reaching the criterion of two successive errorless days on the original task failed to reach the .10 level of significance, either on the B-G or G-W stimulus combinations separately or on the two combined.

Discussion

The results of the comparisons between Groups A, B, C, and D on the test trials strongly suggest that response tendencies to cue-position patterns can be acquired by rats in a situation where the discrimination can be learned on the basis of absolute components alone. While this is contrary to Spence's expectation, to the extent that his statements concerning patterning are interpreted as reflecting necessary conditions, it should be noted that the results fit nicely with the theory if patterning is introduced into the analysis. Thus, Group D would be expected to perform very well on the test problem because S_{BL} and S_{WR} had been consistently reinforced on Task 1, and S_{BR} and S_{WL} consistently nonreinforced on Task 5 during training. Group A with S_{BL} and S_{WR} reinforced during training on Task 1 should also show an initial advantage on the test but to a lesser degree than Group D because S_{BR} and S_{WL} were never presented to Group A during training. Group C with all four patterns reinforced during training on Task 1 should perform in the neighborhood of 50% correct while the transfer of approaching response tendencies to the stimulus patterns, S_{BL} and S_{WR}, by Group B should yield less than 50% correct choices. The obtained percentages of 94, 69, 48, and 36 for the four groups indicate marked agreement with these expectations. At the same time the performance of Groups C and D on the transfer learning task reflects the operation of the response tendencies to the absolute components.

A further implication of the present study concerns the explanation of the faster learning of the so-called simultaneous problem over the successive problem where *S* approaches one of the discriminanda (10, 12). In a Black-White simultaneous problem an analysis including patterned stimuli shows the differential approaching and not approaching tendencies to S_B and S_W to be facilitated by differential approaching and not approaching tendencies to patterns of S_{BR}, S_{BL}, S_{WR}, and S_{WL}. In contrast, the successive problem provides only patterned cues as the basis of the discrimination. It seems possible that the faster learning in the simultaneous problem may be attributable, at least in part, to the larger number of cues available for discrimination in this situation.

The results also raise the question of possible other patterning of stimuli (e.g., transverse patterning) in simple discrimination learning problems. Extrapolation of the present evidence would tend to indicate that the performance of animals in a discrimination task is determined not by approaching and not approaching tendencies either to absolute or to patterned stimuli alone but rather by a function of weighted tendencies to a variety of stimulus elements, both absolute and patterned. The relative weights of the tendencies may well be expected to change as a function of changes in the stimulus situation.

The adequacy of accounting for the data in terms of responses to whole configurations or to relations between stimuli may be examined. One suitable possibility for the original learning might be: "If the whole is darker, choose the left; if the whole is lighter, choose the right." This is comparable to Bitterman's "configurational" patterning (12) (after the analysis of Gulliksen and Wolfle [4]), where within-pairs differentiation has not developed and *S*s respond to the two stimuli as a "whole." While the transfer results of Groups A and B could be accounted for in terms of such an analysis, this is not true of the performance of Groups C and D on the transfer learning tasks where the brightness of the "wholes" was not altered but the

positions of the components were interchanged.

Analysis in terms of within-pairs differentiation (Bitterman's second level of relational learning) encounters difficulties in explaining the learning of the original task. If learning occurred solely on the basis of within-pairs differentiation, one would probably predict that the problem should be more difficult than actually found in this study, and also that many *S*s would fail to meet criterion. Consistent responding to either the darker or the lighter of the two components would appear as position responses to one or the other side and would be rewarded only one-half of the time. Even if *S*s did finally learn the original task on some sort of relational basis, as approach to the extremes of a brightness gradient, such learning would be of no aid in Tasks 2, 3, and 6 where both stimuli presented were either black or white.

Conceivably, the results of this experiment could be accounted for by assuming that both types of differentiation, between-pairs and within-pairs, were developed in the original learning task and that appropriate selection occurred in the transfer tasks. From Bitterman's emphasis (**11, 12**) on the "qualitative distinctiveness" of the two levels of organization and from his labeling of one (configurational) as more "primitive" than the other (relational), one is led to believe that Bitterman restricts each to a particular kind of problem, the successive and simultaneous, respectively. In more recent articles, however, Bitterman (**1, 2**) appears to emphasize these restrictions less and to permit the possibility of the development of different levels of organization in one and the same problem. The latter position would be more similar to the interpretation stated previously regarding the weighted tendencies developed in a discrimination problem.

Summary

Four groups of seven rats each were trained on a brightness discrimination problem which could be learned on the basis of absolute components alone, of patterned stimuli alone, or of both. A transfer learning task presented to two of the groups demonstrates the approaching and not approaching tendencies acquired to absolute components while a test task given to all four groups indicates the presence of response tendencies to patterned stimuli. The latter result is interpreted to be in disagreement with the conditions required for patterning as stated by Spence.

Analysis of the problem including response tendencies to both absolute and patterned stimuli as discrete stimulus elements is seen to account for the results while no simple configurational or relational analysis appears adequate to encompass all the data.

References

1. Bitterman, M. E. Spence on the problem of patterning. *Psychol. Rev.*, 1953, **60**, 123–126.
2. Bitterman, M. E., & Wodinsky, J. Simultaneous and successive discrimination. *Psychol. Rev.*, 1953, **60**, 371–376.
3. Gellerman, L. W. Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, 1933b, **42**, 207–208.
4. Gulliksen, H., & Wolfle, D. L. A theory of learning and transfer. I. *Psychometrika*, 1938, **3**, 127–149.
5. Gulliksen, H., & Wolfle, D. L. A theory of learning and transfer. II. *Psychometrika*, 1938, **3**, 225–251.
6. Krechevsky, I. 'Hypotheses' in rats. *Psychol. Rev.*, 1932, **39**, 516–532.
7. Lashley, K. S. An examination of the 'continuity theory' as applied to discrimination learning. *J. gen. Psychol.*, 1942, **26**, 241–265.
8. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
9. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
10. Spence, K. W. The nature of response in discrimination learning. *Psychol. Rev.*, 1952, **59**, 89–93.
11. Turbeville, J. R., Calvin, A. D., & Bitterman, M. E. Configurational and relational learning in the rat. *Amer. J. Psychol.*, 1952, **65**, 424–433.
12. Weise, P., & Bitterman, M. E. Response selection in discrimination learning. *Psychol. Rev.*, 1951, **58**, 185–195.

(Received January 15, 1955)

ZEITSCHRIFT FÜR
EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

This journal is devoted to experimental investigations covering the entire field of psychology. In addition, up-to-date reviews of current books on psychology are published.

For many years, a strictly scientific journal of psychology has been lacking in German-speaking countries. In taking over this task the "Zeitschrift für experimentelle und angewandte Psychologie" endeavours to continue the tradition of the former, and now extinct, German journals of scientific psychology.

All articles are provided with summaries in English and French.

Herausgeber für Deutschland
Prof. Dr. J. von Allesch, Göttingen; Prof. Dr. Ph. Lersch, München

Herausgeber für Österreich
Prof. Dr. H. Rohracher, Wien

Herausgeber für die Schweiz
Prof. Dr. R. Meili, Bern

Redaktion
Dr. C. J. Hogrefe, Göttingen

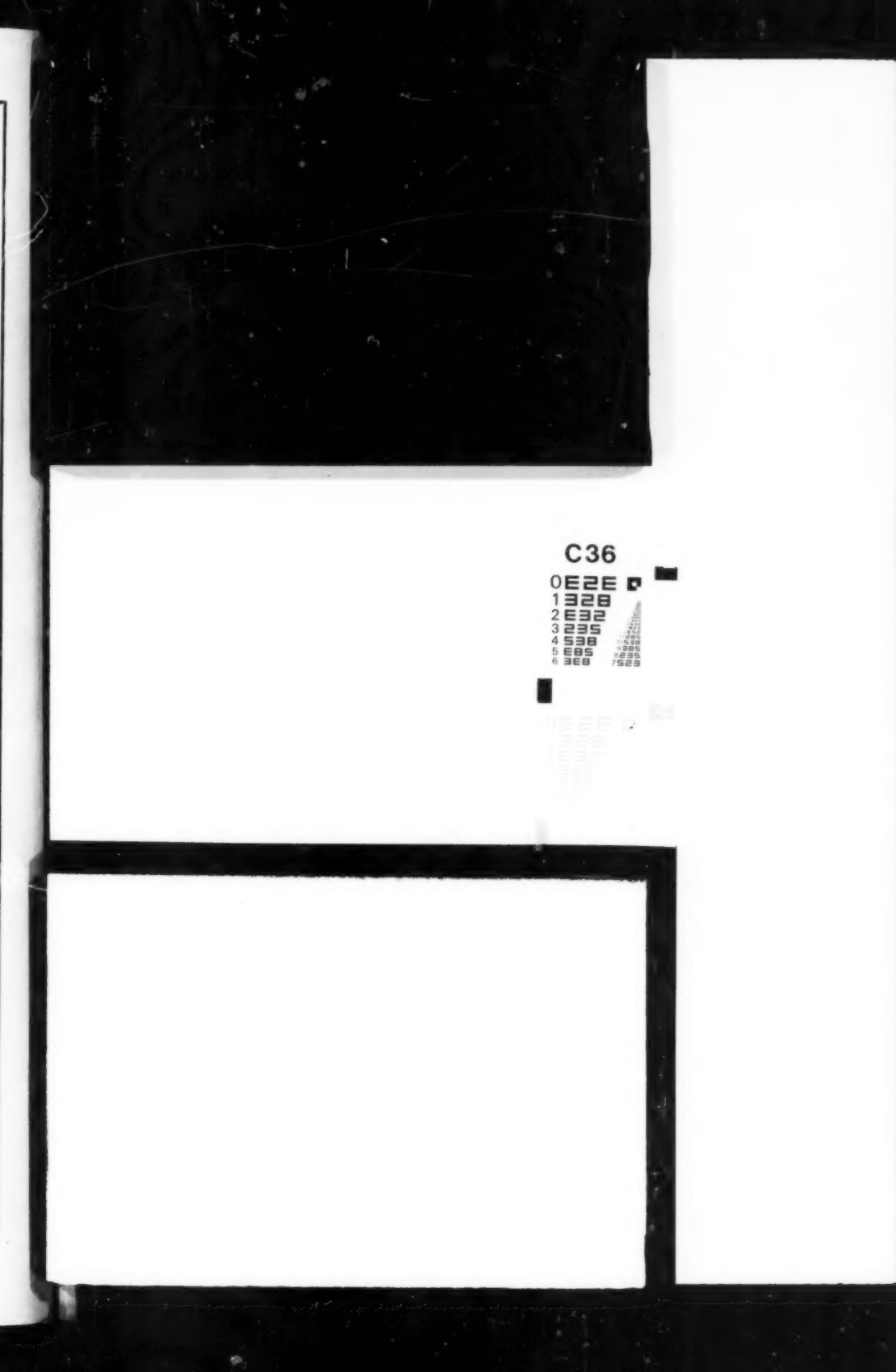
Ständige Mitarbeiter

Prof. Dr. W. Arnold, Nürnberg; Prof. Dr. H. Biäsch, Zürich; Prof. Dr. C. Bondy, Hamburg; Prof. Dr. E. Brunswik, Berkeley USA; Prof. Dr. H. Düker, Marburg; Prof. Dr. G. Ekman, Stockholm; Prof. Dr. Th. Erismann, Innsbruck; Prof. Dr. H. J. Eysenck, London; Prof. Dr. R. Heiß, Freiburg i.B.; Prof. Dr. B. Herwig, Braunschweig; Prof. Dr. H. Hetzer, Marburg; Prof. Dr. E. Kretschmer, Tübingen; Prof. Dr. O. Kroh, Berlin; Prof. Dr. W. Metzger, Münster; Prof. Dr. P. Pichot, Paris; Prof. Dr. E. Rausch, Frankfurt; Prof. Dr. G. Révész, Amsterdam; Prof. Dr. E. Sander, Berlin; Prof. Dr. U. Undeutsch, Köln; Prof. Dr. A. Wellek, Mainz; Prof. Dr. K. Wilde, Göttingen.

The journal appears quarterly and contains 160 pages.

Subscription is DM 64.— per volume and DM 18.— for single numbers.

VERLAG FÜR PSYCHOLOGIE · DR. C. J. HOGREFE · GÖTTINGEN
(Postfach 414, Germany)


C36
0E2E
1328
2E32
3235
4538
5E85
6 3E8